Scuola Internazionale Superiore di Studi Avanzati - Trieste

SISSA 40!

Doctoral Thesis

Topological sorting and self-assembly of knotted molecules: models and simulations

Author:
Mattia Marenda

Supervisor:
Prof. Cristian Micheletti

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

PhD course in Physics and Chemistry of Biological Systems
Molecular and Statistical Biophysics Group

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

Scuola Internazionale Superiore di
Studi Avanzati

Doctoral Thesis

Topological sorting and self-assembly of knotted molecules: models and simulations

Author:
Mattia Marenda

Supervisor:
Prof. Cristian Micheletti

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

PhD course in Physics and Chemistry of Biological Systems
Molecular and Statistical Biophysics Group

November 3, 2018

"I do not know what I may appear to the world; but to myself, I seem to have been only like a boy playing on the seashore, and diverting myself now and then in finding a smoother pebble or prettier shell than ordinary, while the great ocean of truth lay all undiscovered before me."

Sir Isaac Newton (1643-1727)

Acknowledgements

My 4-years PhD journey is about to end. Many people have been part of my life during these wonderful years.
First of all, I would like to thank my supervisor, Cristian. He has been a wonderful guide during this 4-years journey. I have never found a person with such a honesty and such a care for the future of his students. I will never forget his teachings and I will always be thankful to him.
For the first time in my life, I have been part of a work group. I think that I could not have desired better colleagues than you, Stefano, Lucia and Giulia. Thanks for all the brainstormings, chats, and especially for our special "group eatings".
A particular thank goes to Antonio. You have been a wonderful friend and officemate, but also a big brother that helped me growing scientifically during my first years in Trieste.
Trieste has also been the place for me to meet a lot of new friends. I want to thank all the SBP group, past and present members, the directive boards of SISSA Club, the SISSA volleyball team, and every single person that everyday made me feel like home. I will never forget the dinners with Francesca, Matteo and Silvia, the moments of hate and love with Andrea, the confidential sharings with Lorenzo.
Thanks also to the friends that I know since my youth, for having always been by my side, beyond spatial distances. Thanks to Tommaso, Nicolo', Elena, Andrea and Stefano.
Thanks to my family, especially to my mum and my sister, for having always believed in me and for having always supported me. I would not be here without your love.
Thanks to my beloved Giulia, for everything that you do for me. You are the person that filled my heart and I feel very lucky to be by your side. I love you.
My 4-years PhD journey is about to end, but a piece of my heart will always remain in Trieste.

Contents

Acknowledgements **v**

Introduction **viii**

1 Basic concepts on knots **5**
1.1 Introduction to knot theory . . . 5
1.1.1 Definition of knot . . . 5
1.1.2 Classification of prime knots . . . 7
1.1.3 Composite knots . . . 8
1.1.4 Alternating and non-alternating knots . . . 8
1.1.5 Chiral and amphichiral knots . . . 9
1.1.6 Torus knots . . . 10
1.1.7 Twist knots . . . 11
1.1.8 Links . . . 11
1.2 Identifying knot types in closed chains . . . 13
1.2.1 The Alexander Polynomial . . . 13
1.2.2 The Thistelthwaite-Dowker Code . . . 15
1.3 Locating knots in a closed chain . . . 16
1.4 The braid theory and its relation to knots . . . 17

2 Coarse grained models and molecular dynamics simulations **21**
2.1 Coarse-grained model for flexible polymers . . . 21
2.2 Langevin dynamics . . . 22
2.3 Simulation units and time mapping . . . 23

3 Sorting ring polymers by knot type using modulated nanochannels **25**
3.1 Background and motivation . . . 25
3.2 Model and methods . . . 27
3.2.1 The model . . . 27
3.2.2 Molecular dynamics simulations . . . 28
3.2.3 Observables . . . 28
3.2.4 Setting channel parameters . . . 29
3.3 Diffusion in spatially modulated channel . . . 29
3.4 Diffusion of unknotted and knotted rings . . . 31
3.5 The oscillation of $\boldsymbol{D}$ in unknotted chains . . . 32
3.6 The effects of knots . . . 35
3.7 The Fick-Jacobs approximation . . . 37
3.7.1 Fick-Jacobs theory . . . 38
3.7.2 Validity of Fick-Jacobs theory . . . 40

3.7.3 Application to polymers diffusing in modulated channels . . . 40
3.8 Optimal sorting condition for a given chain length . . . 42
3.9 Sorting in presence of a driving force . . . 44
3.10 Summary and perspectives . . . 45

4 Discovering privileged topologies with self-assembly models 49
4.1 Background and motivation . . . 49
4.2 The helical building blocks . . . 52
4.3 Monte Carlo exploration of addressable topologies . . . 53
4.3.1 Monte Carlo method . . . 53
4.3.2 Knot analysis . . . 55
4.3.3 Monte Carlo results . . . 55
4.4 Kinetic accessibility of self-assembling knotted structures . . . 60
4.4.1 Molecular dynamics method . . . 60
4.4.2 Parametrisation of the attractive patchy interaction . . . 62
4.4.3 Parametrisation of the interactions between templates and coordinating particles . . . 65
4.4.4 Molecular dynamics results . . . 66
4.5 Enumerative survey of cyclic-symmetric entangled structures . . . 67
4.6 Summary and perspectives . . . 72

A All single-component knots from the exhaustive braid enumeration 75

To my father

Introduction

Knots are ubiquitous objects and decorative elements that have been studied since antiquity. During the centuries knots have become important not only for their mysterious and elegant aspects, but also for their practical relevance. Knots in ropes, for example, have always been useful for different practical applications, from climbing to sailing, from fishing to medicine.

Chains that are sufficiently long or compactified are prone to develop knots. This is a "statistical necessity" that has been conjectured by Delbruck in 1962 [1] and mathematically proved by Sumners and Whittington nearly 30 years later [2]. In particular, they showed that for a self-avoiding polygon, the knotting probability tends to unity as the polygon length tends to infinity.
This statistical necessity makes topological entanglement a genuine characteristic of polymeric systems. In case of linear polymer chains, knots can be untied by a suitable reptation of the polymer in space and therefore the entanglement is referred as physical knots. On the other hand, if the polymer ends are joined by a cyclisation reaction, the geometrical self-entanglement becomes trapped in the form of a proper mathematical knot, whose topology cannot be changed by any geometrical rearrangement of the polymer except by cutting it [3].
Among polymers, double-stranded DNA (dsDNA) provides an ideal system to study the spontaneous occurrence of knots. In fact, differently from proteins and RNA, metric and topological properties of dsDNA are well captured by aspecific polymer models where only the polymer contour length, persistence length and thickness come into play [4,5]. Studying knots in dsDNA is informative also to understand their biological implication. The presence of knots, in fact, severely affects several cellular processes, such as transcription and replication, with detrimental effects [6–11]. Fortunately, cellular mechanisms have adopted countermeasures: there exist enzymes, namely topoisomerases, that are capable of simplifying the topological complexity of the DNA entanglement by favouring the selective cross-passage of pairs of DNA strands [12–17].
The action of topoisomerases has been understood thanks to the topological profiling of DNA molecules realised with gel electrophoresis. This is the typical technique [18, 19] that permits to sort short DNA molecules by knot type. In particular, molecules are electrically driven through the obstacles of an agarose gel, where their mobility depends on the specific knot type. However, this technique can be used to profile only relatively short DNA molecules ($< 10-15$ kb) [20]. For longer ones, gel electrophoresis resolution would severely degrades, especially for knots with high number of crossings.
This raises the problem of developing novel techniques that can be applied to characterise knot types in longer DNA molecules.
Here, we will use molecular dynamics simulations and theoretical approaches

to discuss the possibility to use spatially modulated nanochannels to sort ring polymer by their knot type [21]. This approach permits, in principle, to separate polymers by their topological complexity, overcoming the aforementioned limits of gel electrophoresis.

The spontaneous knotting of DNA is largely controlled by events where, for example, a loop is threaded by one termini; as a result both the complexity and size of the knots, as well as their location along the DNA contour, are stochastic. This is not the case for other types of biomolecules, particularly proteins, where the folding process towards the native state is tightly controlled by their chemical composition (primary sequence) via their intra-molecular interactions. As a result, proteins whose native state is knotted always feature the same knot type in the same sequence location [22–25] .
Mimicking such reproducible molecular knotting processes are, at least in part, the motivation of the ongoing quest of synthetic chemistry to create synthetic molecules tied in specific knot types [26, 27].
In this regard, chemists succeeded in controlling chemical reactions between small building blocks to assemble molecules with *a priori* desired topology. The chemists who developed this set of techniques, whose contribution opened up the way to a revolutionary chemistry, were awarded with the Chemistry Nobel Prize in 2016 [28].
Despite the high interest in the topic, up to recently, only a handful of different knot types have been synthesised. The reason is due to various challenging aspects of the synthesis process. These include the choice of the suitable building blocks, their correct spatial arrangement, and, above all, the selection of the designable target topology. Not every knot type, in fact, is necessarily expected to be equally designable in practice [29].
In this thesis, we performed a computational and theoretical study to explore which designable molecular knots could be accessible for molecular synthesis with current experimental techniques [30].

The content of this thesis is organised as follows.

In Chapter 1, I will provide a brief introduction of knot theory which sets a reference for concepts used in the subsequent chapters and I will present the methods that we will use to detect knot types and to localise knotted portions in closed chains. At the end of the Chapter, I will briefly present the relationship between braid and knot theory.

In Chapter 2, I will introduce the coarse grained-model of the full flexible polymers used in the following chapter and the Langevin dynamics used in our molecular dynamics simulations.

In Chapter 3, I will focus on the problem of sorting ring polymers by their knot type using their diffusion properties inside spatially modulated nanochannels. I will explore the complex interplay between channel geometry, chain length and chain topology, focusing on the role of knots in the diffusive process. A suitable

interplay between chain and channel parameters, will show that knot sorting is indeed possible and we will see that it holds also in presence of a suitable driving force. This study would serve as a proof of concept for the experimental realisation of channels with optimal sorting capabilities for different topologies.

In Chapter 4, I will present a theoretical and computational survey exploring which new designable molecular knots could be proposed for experimental realisation. The molecular knots that have been assembled so far, show two common features in their geometrical representation, namely quasi-planarity and cyclic symmetry. In this Chapter, I will use self-assembly simulations of rigid helical templates to explore which designable topologies with the above mentioned features could be assembled. In particular, I will use both Monte Carlo and molecular dynamics simulations. The designable topologies emerging from these studies, will include all known molecular knots, as well as new possible targets for future synthesis. At the end of the chapter, by investigating the simple architecture of the resulting constructs, I will present an exhaustive table that recapitulates our results.

The material presented in Chapter 3 and 4 is mainly based on the following published papers:

- M. Marenda, E. Orlandini and C. Micheletti
"Sorting ring polymers by knot type with modulated nanochannels"
Soft Matter ***10****, 795, (2017)*

- M. Marenda, E. Orlandini and C. Micheletti
"Discovering privileged topologies of molecular knots with self-assembling models"
Nature Communications ***9****, 3051, (2018)*

A collaborative project is also currently ongoing to introduce hydrodynamics effects in the sorting investigation of Chapter 3:

- L. B. Weiss, M. Marenda, C. Micheletti and C.N. Likos
"Translocation of polymers of various knot types through modulated channels with MPC"
Expected submission date: December 2018

Chapter 1

Basic concepts on knots

In this Chapter, I will present a brief introduction of knot theory, focusing on concepts that I will use in the following chapters [31, 32]. After defining knots and their main properties, I will treat the problems of knot detection and localisation in closed curves [33]. Finally, the relation between braid and knots will be briefly discussed [34].

1.1 Introduction to knot theory

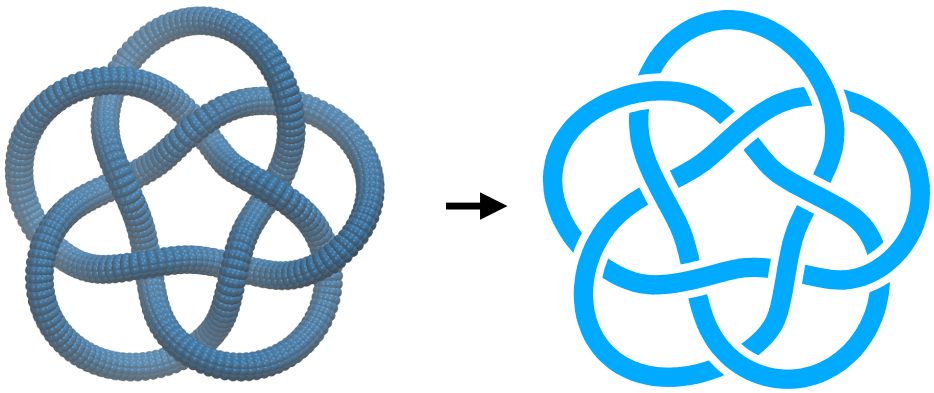


Fig. 1.1: Example of 3-dimensional knot and its projection into a 2-dimensional diagram. Most of the figures of knots of this chapter were made using the Knotplot software [35].

1.1.1 Definition of knot

Everyday we inevitably experience the effects of knots. Just imagine, for example, when you get annoyed while trying to unknot your earphones. Our daily experience can actually help us in defining a *knot*. Suppose to grab a rope, tie a simple knot in it and then glue the rope ends together. We can rearrange it in space as much as we want but, as long as we preserve its structural integrity, its initial knotted topology will always be maintained. The only way to unknot our rope is by cutting and re-opening it.
A *knot* can indeed be defined as the set of continuous simple and differentiable closed curves that are related by ambient isotopies in $\mathbb{R}^3$. These are geometrical manipulations of the curve (isotopies) through the space in which the curve is

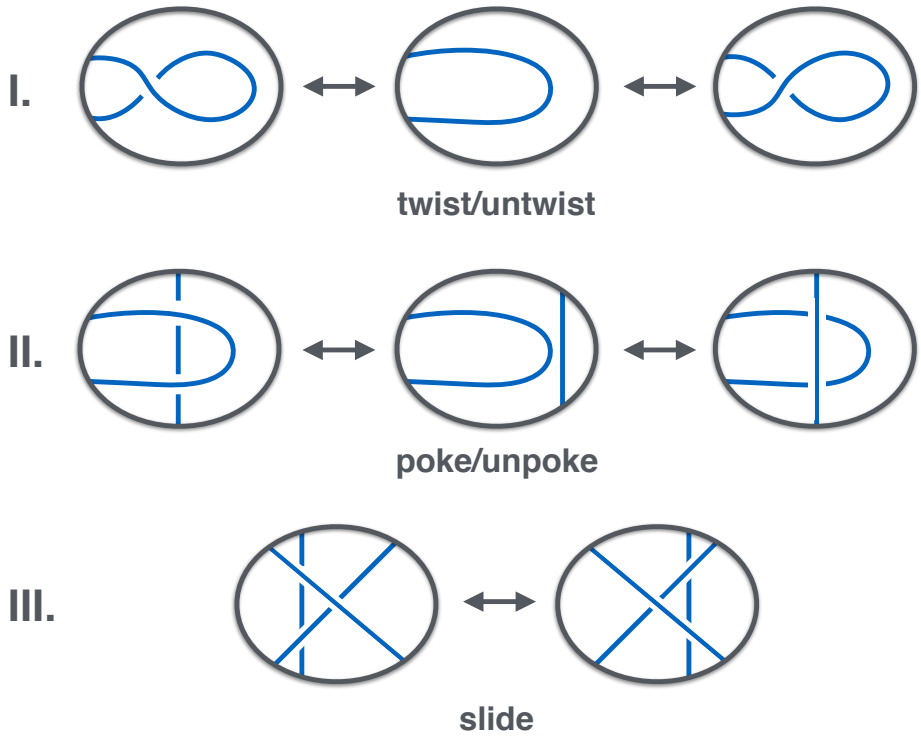


Fig. 1.2: The three possible types of Reidemeister moves.

embedded (ambient). Note that ambient isotopies are transformations that do not allow to shrink the knot down to a point to be rid of it, just as you cannot do it in a string by pulling the knot tighter and tighter.
Instead of considering curves in the 3-dimensional space, one can alternatively look at their planar representations, the so-called *knot diagrams*, where *crossings* are represented by interrupting one of the branches of the projected curves, as shown in Fig. 1.1.
The equivalent of the ambient isotopies that applies to planar diagrams, are the so-called *Reidemeister moves*. These are only three possible types of planar moves that change the projection of the knot, without altering the topology, see Fig 1.2. It has been demonstrated that two different knotted diagrams are equivalent if and only if one can be deformed into the other one via a succession of these moves.

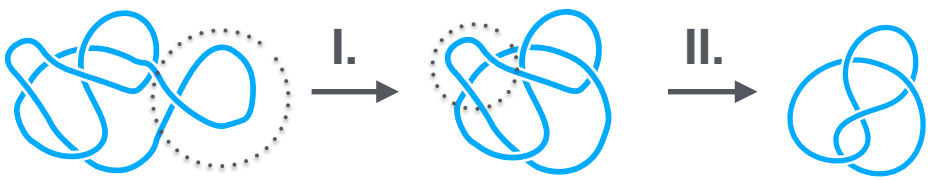


Fig. 1.3: Two consecutive Reidemeister moves are applied to a knot diagram to obtain its minimal representation.

1.1.2 Classification of prime knots

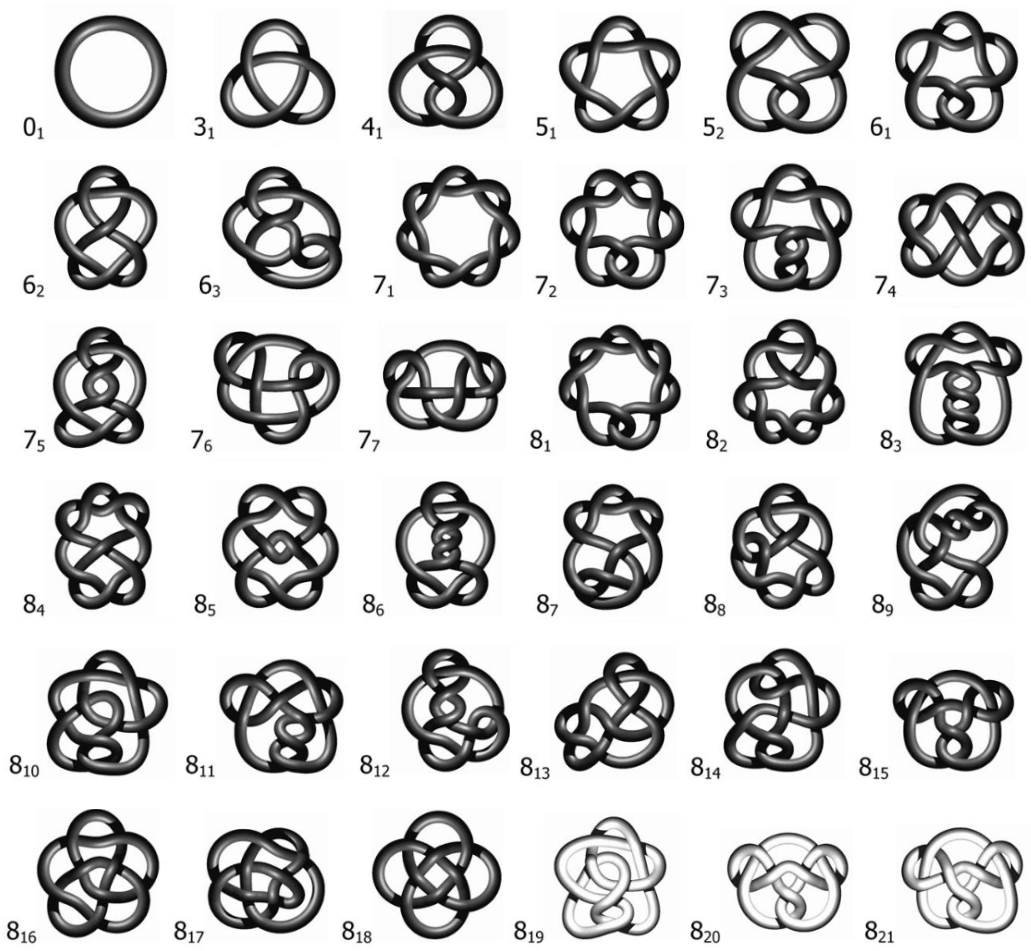


Fig. 1.4: Knots classification up to 8 essential crossings. Knots 8_{19}, 8_{20} and 8_{21} are the simplest instances of non-alternating knots. Taken from [36].

By using Reidemeister moves, one can, in principle, recast any knot diagram into its *minimal representation*, which corresponds to the diagram with the smallest number of crossings C, the so-called *crossing number*, see Fig. 1.3.
As a matter of fact, the value of C provides the most used criterion for classifying *prime knots* (i.e. those that cannot be decomposed as a connected sum of simpler knots) in groups of increasing complexity. The number of distinct topologies for a given crossing number C grows approximately exponentially [37], reaching a total amount of 1.7 million different prime knots for $C \leq 16$ [38].
The notation that we use here for knot classification, was first introduced by Alexander and Briggs in 1926 [39]. They labelled every knot type with C_i, where C is the crossing number, while i is a conventional enumerative index introduced to distinguish different knots with the same nominal complexity.
In Fig. 1.4 we can find the minimal representation of knots with up to eight crossings. The trivial knot, 0_1, is often called the unknot. The simplest non-trivial knot is the trefoil knot, 3_1, which is characterised by three crossings in its

minimal representation. Then, we have the figure-of-eight knot, 4_1, with four crossings. Starting from $C = 5$, we have more than one knot with the same complexity and hence the enumerative index is essential to distinguish them. In particular we have two knots with $C = 5$, the 5_1 and 5_2 knots, three with $C = 6$, seven with $C = 7$, twenty-one with $C = 8$, and so on. Exhaustive knot tables with the Alexander-Briggs notation were published by Rolfsen in 1976 [40] up to $C = 10$.

1.1.3 Composite knots

As common experience teaches us, a non-trivial prime knot C_i in a rope cannot be untied by introducing a second non-trivial knot C_j in a different portion of the same rope, no matter how we manipulates the doubly knotted rope. Essentially, knots obtained in this way are termed *composite* knots and are indicated with $C_i \# C_j \# \ldots$. The prime knots involved in the composition $C_i, C_j, \ldots$ are called *factors* of the composite knot.
Knot composition is commutative and associative. Also, composite knots have a unique and finite factorisation into prime knots.
An example of a composite $3_1 \# 3_1$ knot is shown in Fig. 1.5.

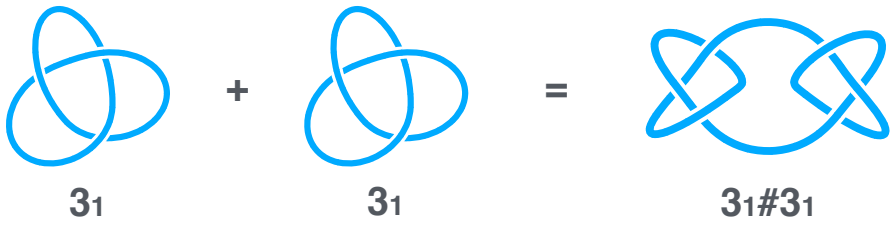


Fig. 1.5: Composition of two prime knots with three essential crossings (3_1).

1.1.4 Alternating and non-alternating knots

Fig. 1.6: The 4_1 knot in its minimal alternating representation and in one of its non-alternating geometries. An arbitrary orientation has been attached to the curves.

In Chapter 4, we will resort the concept of *alternating* and *non-alternating knots*. In *alternating knots*, the crossings in the minimal representation alternate between under and over as one travels along the oriented knotted curve (the choice of the orientation is not important). Interestingly, the minimal representation of knots with $C < 8$ are all alternating, and the simplest non-alternating instances are the 8_{19}, 8_{20} and 8_{21} knots, as shown in Fig. 1.4. Note also that alternating knots do not necessarily maintain this feature in their non-minimal representations. See for example the minimal diagram of the 4_1 knot and one of its non-minimal diagrams that is non alternating in Fig. 1.6.

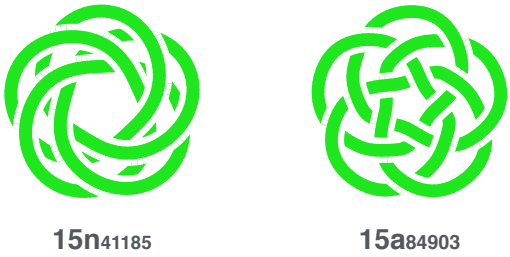


Fig. 1.7: Minimal diagram of an alternating and a non-alternating knot with 15 crossings. Their name in Thistlethwaite notation is specified.

The alternating/non-alternating feature of knots is not explicited in Alexander-Briggs notation, though it is highlighted in the recent nomenclature introduced by Hoste and Thistlethwaite in the 1990s [38]. In particular, this feature is indicated respectively with an a or a n between the number of crossings and a conventional enumerative index, as shown in the label of the knots in Fig. 1.7. Knot tables with the Thistlethwaite notation have been published up to $C = 16$. In this thesis, I will use the standard Alexander-Briggs indexing scheme up to $C = 10$ crossings, and the Thistlethwaite one for more complex knots up to $C = 16$.

1.1.5 Chiral and amphichiral knots

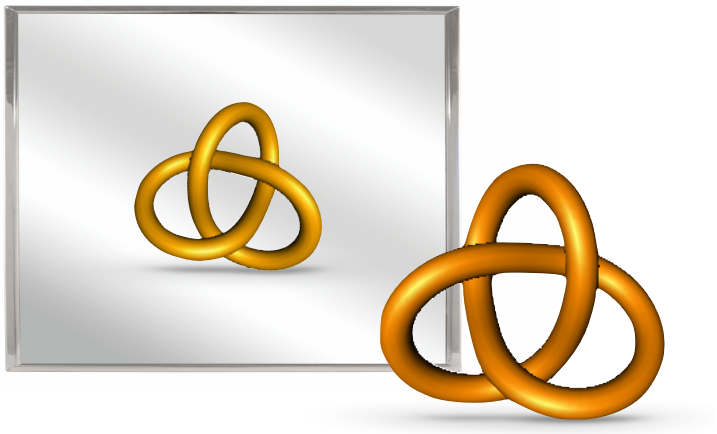


Fig. 1.8: The chiral trefoil knot 3_1 and its mirror image.

Similarly to chiral molecules in chemistry, a knot is said to be *chiral* if its minimal representation cannot be deformed with an ambient isotopy into its mirror image. Otherwise, it is said to be *amphichiral*. The simplest example of chiral knot is the 3_1 knot, shown in Fig. 1.8.
Amphichiral knots are uncommon among simple knot types, in fact among all the knots up to 8 essential crossings, only 1/4 of them is amphichiral, namely the 4_1, 6_3, 8_3, 8_9, 8_{12}, 8_{17} and 8_{18} knots.

1.1.6 Torus knots

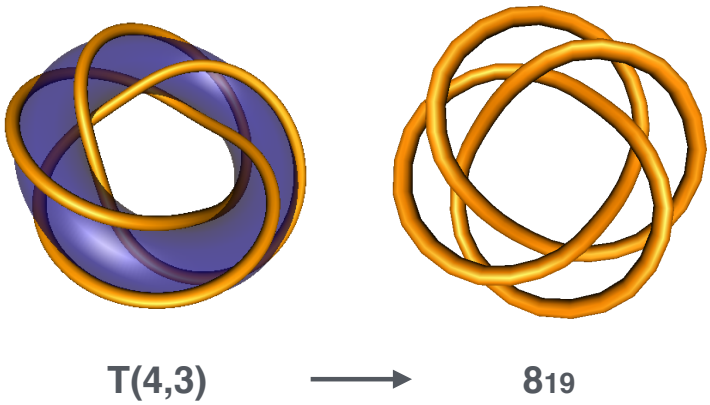


Fig. 1.9: Representation of the torus knot $T(4,3)$, whose geometry topologically represents the 8_{19} knot.

Knots can be grouped in families based on notable or distinctive features of their minimal representation. Two of them will be encountered in Chapter 4: the *torus knots* and the *twist knots*.
The family of *torus knots* includes all prime knots that can be drawn as a simple closed curve on the surface of a torus.
A general parametrisation in $t \in [0, 2\pi]$ of torus knots is:

$$\begin{cases} x = -\sin(qt)\,[R + r\cos(pt)] \\ y = \cos(qt)\,[R + r\cos(pt)] \\ z = r\sin(pt) \end{cases} \tag{1.1}$$

where the parameters r and R control the geometry of the torus: r is the radius of the tube and R radius of the hole. The p and q numbers are coprime integers that can be either positive or negative, depending on the orientation of the winding direction, and completely classify the (p, q)-torus knots, also denoted as $T(p, q)$. Geometrically, p represents the number of times that the knot meridionally winds around the torus, while q the number of times it winds equatorially. Note that if p and q are not mutually prime, the definition can be extended to a multi-component torus knot, or torus link (see next sections for examples).
The simplest non-trivial torus knots, in order of nominal complexity, are the 3_1, 5_1, 7_1, 8_{19}, 9_1, 10_{124} knots and they are obtained with different combinations

of p and q. An example of a torus knot is shown in Fig. 1.9, where $p = 4$ and $q = 3$. This geometry topologically represents the 8_{19} knot.
Key properties of torus knots are:

- A (p, q)-torus knot is topologically equivalent to the (q, p)-torus knot.
- If p or q is 1 or -1 the curve is unknotted and can be drawn on the surface of a sphere.
- All non-trivial torus knots are chiral.
- The crossing number of a torus knot is given by $\min\{p(q-1), q(p-1)\}$.
- Every torus knot $T(p, q)$ is symmetric for discrete rotations around the torus axis. The symmetry group is C_p (therefore the symmetry is by rotations of an angle $\gamma = 2\pi/p$.)

1.1.7 Twist knots

The family of *twist knots* includes all knots that can be obtained by repeatedly twisting an unknotted ring and then clasping the two ends together, as shown in Fig. 1.10. The simplest twist knots, obtainable with an increasing number of half-twists, are the 3_1, 4_1, 5_2, 6_1, 7_2 and 8_1 knots. Twist knots can be chiral or amphichiral and their crossing number is equal to $n + 2$, where n is the number of half-twists. Note that the trefoil knot belongs both to the family of torus and twist knots and it is also an alternating knot.

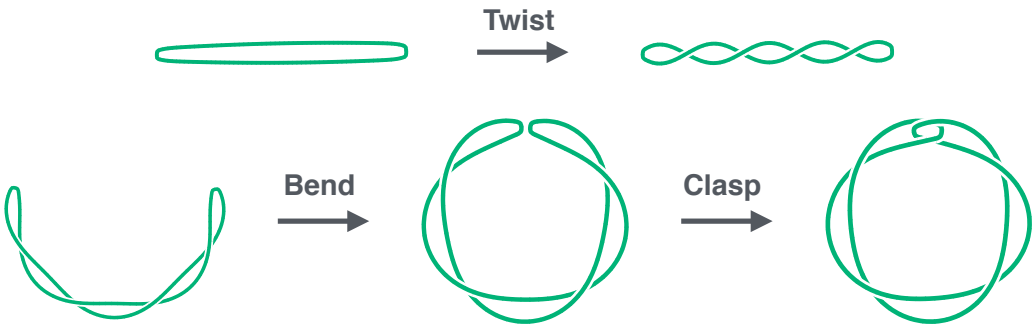


Fig. 1.10: Formation of a twist knot with four half-twists. This geometry represents the 6_1 knot.

1.1.8 Links

So far, we have restricted our attention to knots; that is to say, single entangled curves. In Chapter 4, we will encounter also two or more closed curves entangled together. These are multi-component knots, that are referred as *links*. Each closed curve, that can be either knotted or not, is called a *component* of the link and the number of components is called *multiplicity* (m).
One can define also for links equivalence classes that include all configurations

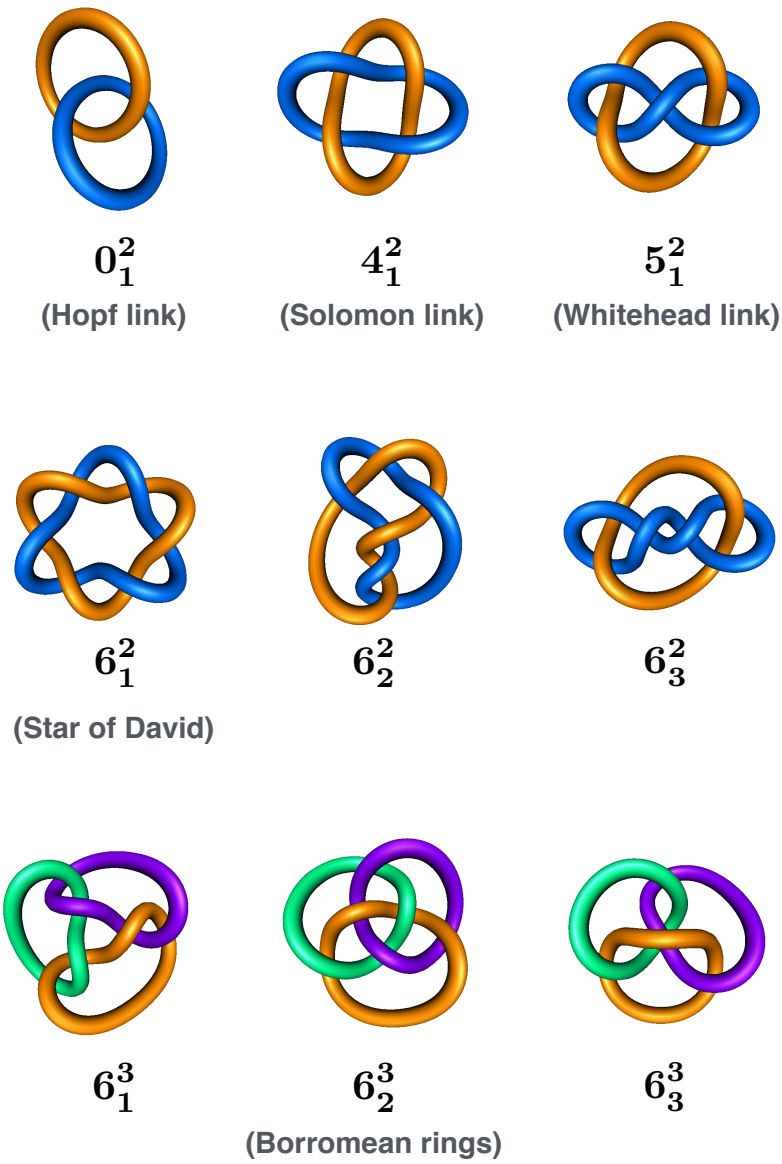


Fig. 1.11: Links with up to 6 crossings of their minimal diagram. Multiplicity is either 2 or 3.

obtainable with an ambient isotopy (or a generalised set of multi-component Reidemeister moves for planar diagrams).
Similarly to knots, links are customarily tabulated for increasing nominal complexity of their minimal diagrammatic representation and therefore by the value of the crossing number C.
The commonly employed notation for links specifies both the crossing number and the multiplicity of the link, the latter as a superscript, see Fig. 1.11. Furthermore, an enumerative subscript index is added to distinguish links with the same nominal complexity and the same multiplicity, but different topology.
In Fig. 1.11 link types with up to 6 crossings are shown. Note that these consists in 2 or 3 unknotted curves intertwined in different ways.
The 4_1^2 and 6_1^2 links are also called Solomon link and Star of David, after their

cultural and symbolic use. They are the simplest examples of torus links, that emerge when p and q are not coprime. Other simple link types include the Borromean rings, e.g. the 6_2^3 link of Fig. 1.11, that consists of three interlocked curves that are pairwise unlinked.

1.2 Identifying knot types in closed chains

1.2.1 The Alexander Polynomial

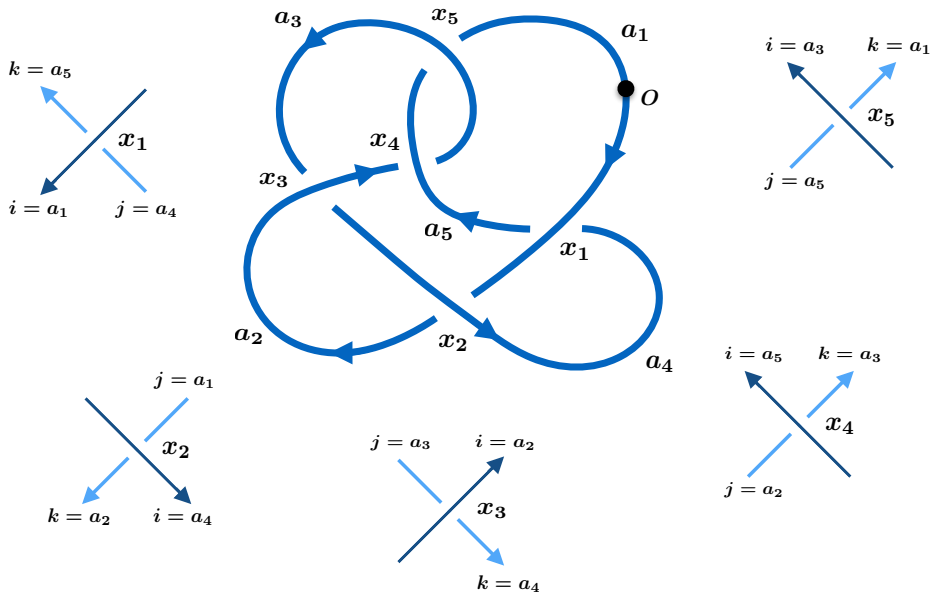
Fig. 1.12: Positive and negative crossings defined with the right-hand rule. The arc labelling i, j, and k is for the computation of the Alexander polynomial. Adapted from [3].

In the search of designable topologies of Chapter 4, we will deal with thousands of computer generated knotted curves and we will need to establish their topology. In numerical contexts, a powerful strategy to identify the knot type of a given curve, is to associate a *topological invariant* to the curve. This is a quantity that depends only on the knot type and not on the details of its geometrical representation. In particular, as topological invariant, we shall mainly use the *Alexander Polynomial* [41] due to its simplicity and its computational efficiency (for relatively simple geometrical representations). This polynomial is defined in terms of a single variable t and is computed starting from a given diagram according with the following algorithm [3]:

1. Assign an orientation to the knot diagram and establish the sign of each crossing following the right-hand rule as shown in Fig. 1.12.

2. Starting from an arbitrary non-crossing point, O, on the curve, follow the curve orientation and assign a progressive numbering index to the n crossings and to the n arcs between them.

3. Define an $(n \times n)$ matrix M. The rows of M correspond to the crossings, while the columns to the arcs. The elements of row x are calculated by considering the x^{th} crossing in the diagram and the three arcs i, j, and k taking part to the crossing. Without loss of generality, we will assume that the arc i passes over arcs j and k. All elements of the the row are set to zero except for $M(x,i)$, $M(x,j)$ and $M(x,k)$. These are calculated as follows:



$$M(t) = \begin{pmatrix} 1-t & 0 & 0 & t & -1 \\ t & -1 & 0 & 1-t & 0 \\ 0 & 1-t & t & -1 & 0 \\ 0 & t & -1 & 0 & 1-t \\ -1 & 0 & 1-t & 0 & t \end{pmatrix}$$

$$\Delta(t) = \begin{vmatrix} 1-t & 0 & 0 & t \\ t & -1 & 0 & 1-t \\ 0 & 1-t & t & -1 \\ 0 & t & -1 & 0 \end{vmatrix} = 2t^3 - 3t^2 + 2t$$

$$\Delta(t; 5_2) = 2t^2 - 3t + 2 \qquad \begin{cases} \Delta(-1, 5_2) = 7 \\ \Delta(-2, 5_2) = 16 \to 1 \end{cases}$$

Fig. 1.13: Computation of the Alexander polynomial for the 5_2 knot. Adapted from [3].

- If x is a positive crossing, then $M(x, i) = 1 - t$, $M(x, j) = -1$ and $M(x, k) = t$.
- If x is a negative crossing, then $M(x, i) = 1 - t$, $M(x, j) = t$ and $M(x, k) = -1$.
- If x is a degenerate crossing emerging form a twist of the curve ($i = k$ or $i = j$), then $M(x, j) = 1$ and $M(x, k) = -1$ irrespective of the crossing's sign.

4. Delete any one of these columns and any one row to obtain an $(n-1) \times (n-1)$ matrix. This is the *Alexander matrix* associated to the diagram. Its determinant is the so-called *Alexander polynomial* $\Delta(t)$.

Strictly speaking, the Alexander polynomial is not uniquely defined for a given knot type, because the size of the matrix and hence the determinant depend on the number of crossings and therefore on the details of a given diagrammatic representation. However, it turns out that the Alexander polynomials of two different diagrammatic representations of the same knot type can differ only by a multiple of $\pm t^m$, $m \in \mathbb{Z}$. By factoring out and removing such $\pm t^m$ contributions, one obtains a polynomial, referred as *irreducible Alexander polynomial* $\Delta(t; C_i)$, that is independent on the geometrical representation of the knot.
As in many computational studies, we will not determine the symbolic Alexander polynomial to identify the knot type, but we rather compute its values in $t = -1$ and $t = -2$, $\Delta(-1; C_i)$ and $\Delta(-2; C_i)$. Note that that these two values depend on the knot representation respectively by a factor $(-1)^m$ and $(-2)^m$ (as a consequence of the $\pm t^m$ dependence mentioned above).
The calculation of the Alexander Polynomial for the 5_2 topology is given in Fig. 1.13.
In our investigations, to identify the knot types of our chains, we used the Kymoknot software [33], where the calculation of Alexander polynomials is efficiently implemented.
Note that, even if the Alexander polynomial is a topological invariant, different knot types are not guaranteed to have different Alexander polynomials. There are, in fact, many examples of distinct topologies with the same polynomial. For instance, the irreducible Alexander polynomial of 8_{20} is $\Delta(t; 8_{20}) = (t^2 - t + 1)^2$, which is equal to that of the composite $3_1 \# 3_1$ knot. This ambiguity can appear also when dealing with complex prime knots. In fact, they could share the same Alexander determinant of simpler knots or even of the unknot.

1.2.2 The Thistelthwaite-Dowker Code

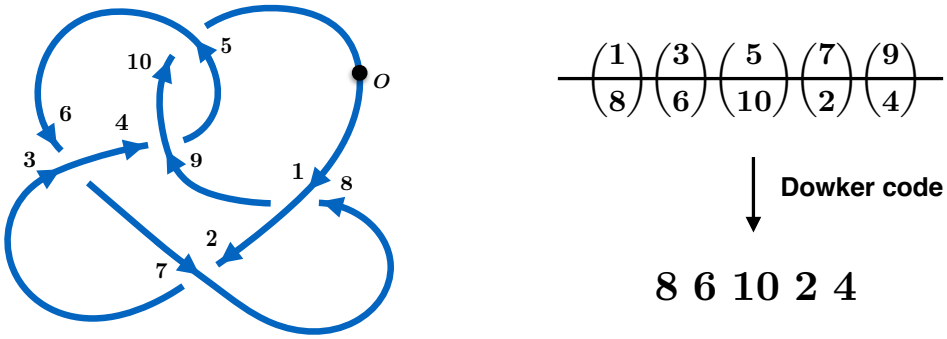


Fig. 1.14: Computation of the Dowker code for the diagram of a 5_2 knot for an arbitrary choice of O and the curve orientation.

In case of complex knot representations, as just mentioned, the knot type identification might be ambiguous when relying only on Alexander polynomials. Furthermore, the calculation of the Alexander polynomial might become quite expensive in terms of computational time in case of complex geometrical representations. The size of the Alexander matrix grows, in fact, cubically with

the number of crossings. To overcome this limitations, as a complementary approach we use the *Thistelthwaite-Dowker code* [38], also simply referred as *Dowker code*, for the identification of complex knot types. The Dowker code is a sequence of signed integers that algebraically encodes for a specific diagram with n crossings and is calculated as follows:

1. Assign an orientation to the knot diagram and an arbitrary starting point O.

2. Follow the curve from the starting point and label each of the n crossings encountered with increasing integers 1, ..., $2n$. Each crossing thus receives two integer labels, one even and one odd, and a one-to-one correspondence is defined between the set of odd labels and the set of even labels.

3. Assign a minus sign to each even integer which is the label of an overpass. The overcrossing-undercrossing structure is therefore captured.

4. Write the pair of numbers in the natural order of odd labels and consider just the sequence of the associated even numbers. This is the *Dowker code* for the given choice of starting point.

An example of Dowker code calculation for the minimal diagram of 5_2 knot is shown in Fig. 1.14.
The Dowker code can often be algebraically simplified to its minimal form and then compared to Thistlethwaite tabulated values to identify the knot type for $C \leq 16$. Actually, the Dowker code simplification to its minimal form is not always guaranteed, and therefore the determination of the knot type. However, it was sufficient to identify the knot type of almost all the curves considered in this thesis.
The calculation of the Dowker code, its simplification and the identification of the knot type were performed with the *Knotscape* software package [42].

1.3 Locating knots in a closed chain

Alongside the identification of the knot type, we are interested also in locating the knotted portion of a given chain. In Chapter 3, the knotted portion location will help understanding the properties of knots inside spatially modulated nanochannels.
The procedure that we follow, involves a bottom-up search for the knot with the purpose of finding the shortest portion of the ring that has the same topology as the entire chain. One starts by considering all portions of the ring with a very small contour length, l (small means no larger than the minimal length required to tie a trefoil knot). If, after a suitable closure (see below), none of these portions has the topology of the entire chain, l is increased and the same search is performed again. The algorithm stops when one portion is found with the target topology.
When applying this knot location scheme, one needs to introduce a well-defined

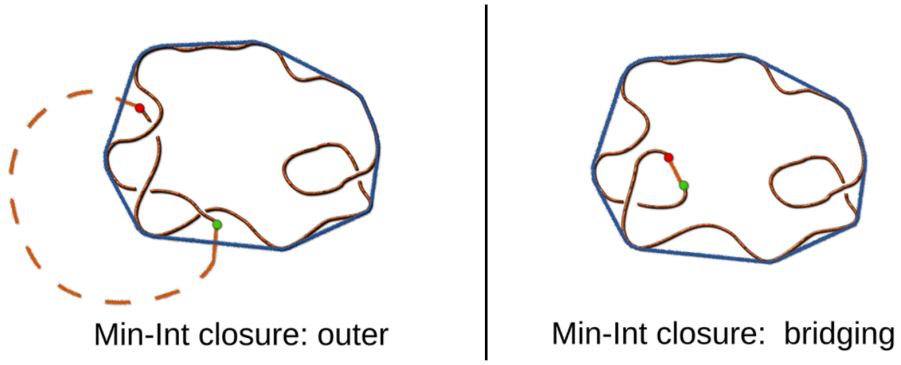


Fig. 1.15: The two possible closure methods in the minimum-interfering closure scheme. This picture is taken from ref. [33].

procedure to close, or circularise, the various subportions of the ring under consideration so as to properly establish their topological state. Different closure schemes have been developed over the years [43–45]. Here, we shall use the so called minimally-interfering closure scheme [44]. This closure method is based on the idea of adding with the closure segment the least amount of entanglement to the considered region. In particular, the open arc is closed in one of the two following ways. The first is directly joining the open ends with a straight segment. The second one is prolonging the two endpoints until the intersection with the convex hull and connecting these intersection points with an arc to infinity. The closure with the smaller total arc length (excluding the closure to infinity) is used. An example of this closure scheme is shown in Fig. 1.15.
The algorithm for knot location used in this thesis has been implemented in the *Kymoknot* software package [33].

1.4 The braid theory and its relation to knots

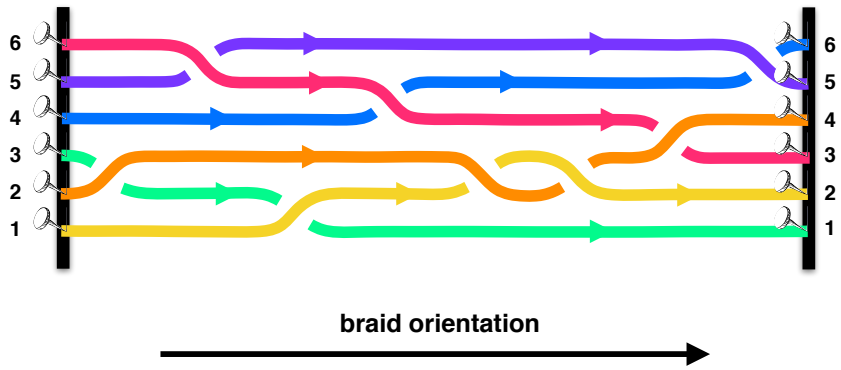


Fig. 1.16: Representation of a braid with $n_S = 6$ strings. The braid orientation is shown both in the strands and at the bottom of the figure. Positions of strands are labelled at the beginning and at the end of the braid.

If we type the word "braid" on Google, the main results are pictures representing several hair styles. All of them have in common a pattern of hair locks intertwined together. Actually, this is probably what we think about when we hear the word braid. More generally, we can think of a *braid* of n_S *strands* as n_S threads attached to vertically aligned nails and horizontally extending rightward while crossing each other. At the end, these threads are also secured to nails. See Fig. 1.16 for an example.
One can easily realise that considering a braid extending rightward is equivalent to assigning an orientation to the braid, as shown in Fig. 1.16. Typically, an index i is used for the strand at i^{th} position (count starts from the bottom). Note that this index is not a static label for a given strand, because it changes along the braid span when strands cross each other. Braid crossings, in fact, define points where two strands exchange their position. As an example, we observe that the orange strand in Fig. 1.16 starts from position $i = 2$, but then it passes to positions 3, 2, 3 and 4 upon crossing other strands.
A specific braid does not have a unique representation, because its strands can be displaced and hence new crossings introduced, while keeping their starting and final points fixed and without introducing breaks. This means that also in braid theory every representation belongs to a specific class of equivalence.
We can define the *composition* or *product* operation on the set of braids with n_S strands. It consists in joining the strands at position i at the end of the first braid with the strands at the same position at the beginning of the second braid, as shown in Fig. 1.17.

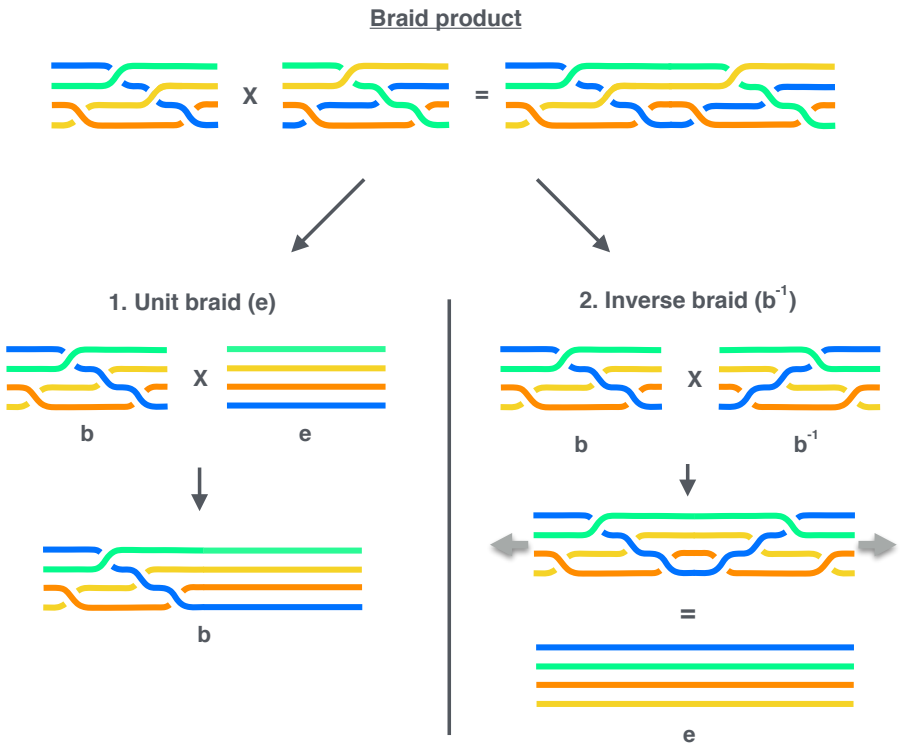


Fig. 1.17: Representation of the products of braids with $n_S = 4$. Unit braid e (left) and inverse braid b^{-1} (right) are shown in the bottom part of the figure.

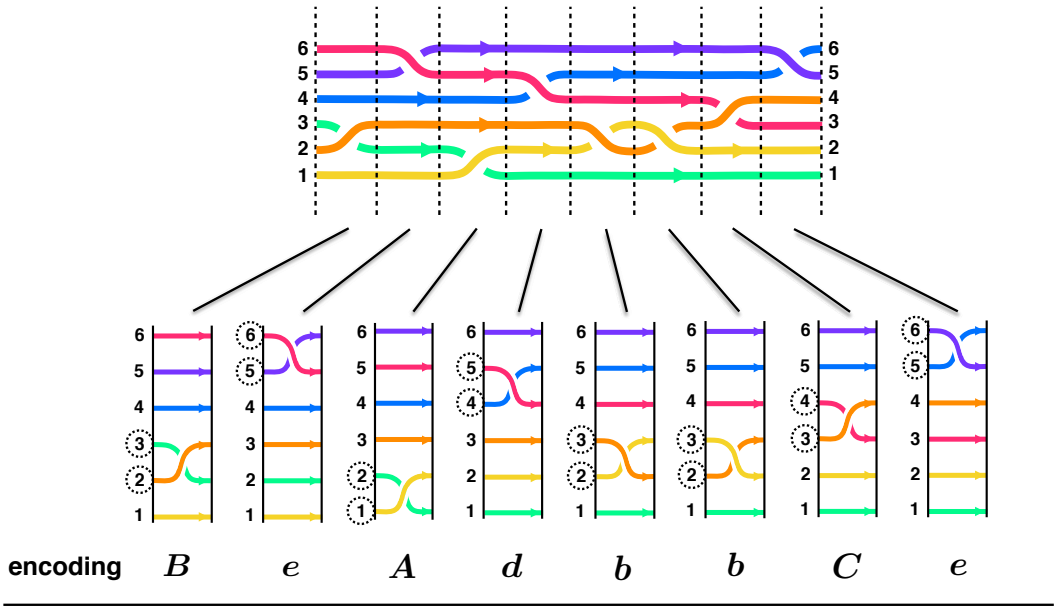


Fig. 1.18: Representation of a braid with $n_S = 6$ strands. It is formed by the product of 8 essential braids. Below, the name and the shortcut that characterise the braid.

The set of braids with n_S strands (with $n_S > 2$) endowed with the above mentioned *product* ($\times$) define a group, denoted by B_{n_S}, which satisfies the group axioms:

- There exists an identity element known as the *unit braid* (e), such as $b \times e = b$. This is the *trivial braid*, whose strands hang horizontally without crossing (see Fig. 1.17.).
- There exists an *inverse braid* (b^{-1}) for every braid b, such as $b \times b^{-1} = e$. The inverse braid is the vertical mirror image of the given braid (see Fig. 1.17).
- The product of braids is *associative*: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

The product of braids makes it possible to replace their graphical representation with a symbolic string or algebraic encoding. In fact, every braid can be decomposed as a product of elementary braids, the so-called *essential braids*. These contains just a single crossing taking place between the two strands at i and $i+1$ position, as shown in Fig. 1.18. Every essential braid can be denoted with the i^{th} alphabetic letter, in lower or uppercase if the crossing is respectively positive or negative (see Fig. 1.12). In this way, one obtains a sequence of letters, or a word, associated to the braid. An example of symbolic encoding is shown in Fig. 1.18.
Similarly to what was discussed for links, one can take the intertwined braid strands, "bending" them and glueing together the two ends of the strands at the same position i, as shown in Fig. 1.19. This is referred to as *braid closure* and the resulting entangled structures are knots or links. The multiplicity of the final linked construct depends on the specific crossing pattern.

Actually, in this way we could, in principle, obtain all possible knots or links as proved by J. W. H. Alexander in 1923 in his celebrated theorem [46].
An example is shown in Fig. 1.19, where the closure of the 3-strands braid *abababababab* results in the 10_{124} topology.

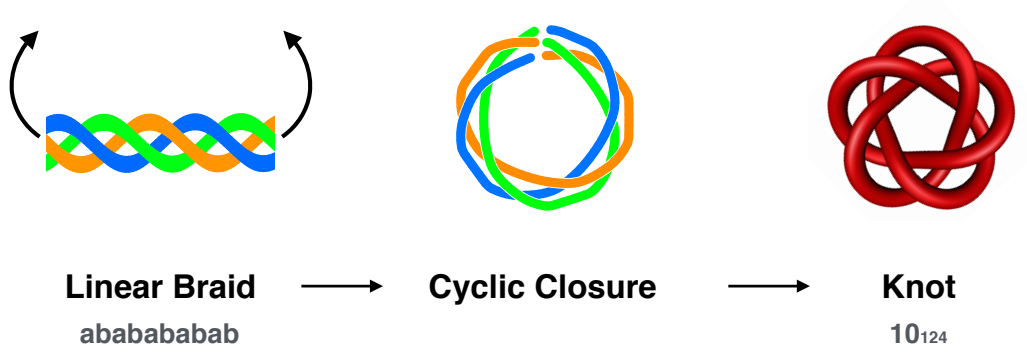


Fig. 1.19: Cyclic closure of the linear braid *abababababab*. As a result, the 10_{124} torus knot is obtained.

Chapter 2

Coarse grained models and molecular dynamics simulations

In this Chapter, I will provide an introduction to the general methodologies used in Chapter 3, covering both models and simulations.

2.1 Coarse-grained model for flexible polymers

In Chapter 3, we will consider circular *homopolymers*, i.e. linear macromolecules that consist of identical elementary units, called monomers, arranged in a circular fashion [47]. We schematically model a fully-flexible ring polymer as a

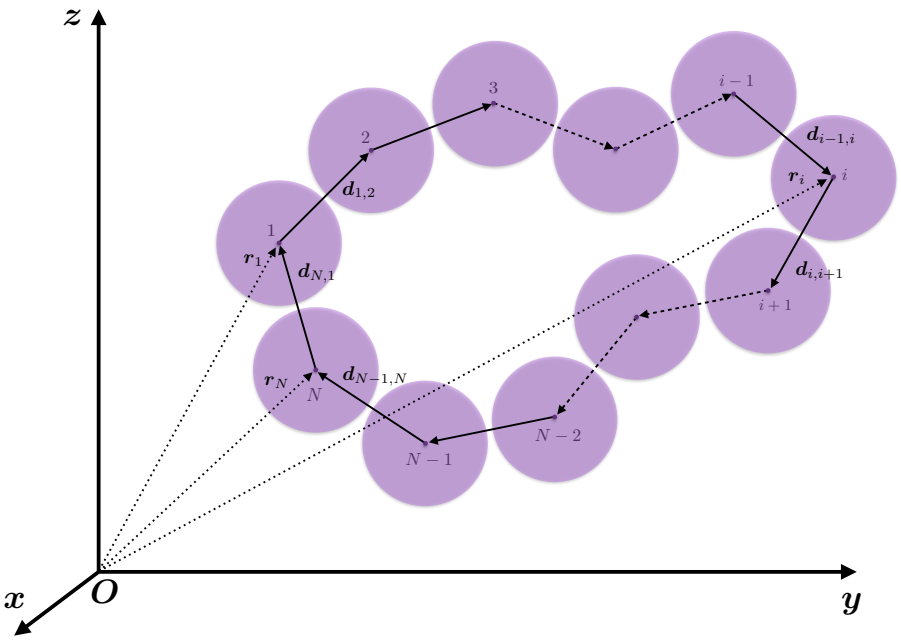


Fig. 2.1: Sketch of a ring polymer modelled as a chain of N equal beads.

chain of beads with each bead representing a monomer. The model is sketched in Fig. 2.1, where the beads positions are indicated by $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, ..., $\boldsymbol{r}_N$, while the distance vectors between beads by $\boldsymbol{d}_{i,j} = \boldsymbol{r}_j - \boldsymbol{r}_i$. The bond vectors, $\boldsymbol{d}_{i,i+1}$, are highlighted in Fig. 2.1.

To account for the volume occupied by each bead, a steric interaction is introduced between every pair of beads i and j. In particular, we used a truncated and shifted Lennard-Jones potential:

$$U_{LJ}(d_{i,j}) = \begin{cases} 4C_{LJ}\epsilon \left[\left(\frac{\sigma}{d_{i,j}}\right)^{12} - \left(\frac{\sigma}{d_{i,j}}\right)^{6} + \frac{1}{4} \right] & \text{if } d_{i,j} < 2^{1/6}\sigma \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

where σ is the nominal bead diameter, ϵ the unit of energy, and C_{LJ} is the magnitude of the potential, usually set to $C_{LJ} = 1$.
Note that U_{LJ} and its derivative are continuous.
An additional potential energy term is introduced to account for chain connectivity, i.e. the bonding of consecutive beads. As customary, we used the so-called *finite extensible non-linear elastic (FENE)* potential [48]:

$$U_{FENE}(d_{i,i+1}) = -\frac{1}{2} C_F \, \epsilon \left(\frac{R_0}{\sigma}\right)^2 \log\left[1 - \left(\frac{d_{i,i+1}}{R_0}\right)^2\right] \tag{2.2}$$

where $C_F = 30$ and $R_0 = 1.5\ \sigma$ are the potential magnitude and the maximum bond length.
The sum of the two potentials, $U_{LJ}(d_{i,i+1}) + U_{FENE}(d_{i,i+1})$, favours an effective bond length of about σ (the bead nominal diameter), see Fig. 2.2. As a consequence, one obtains a chain of beads where the bond crossing is highly unfavoured.

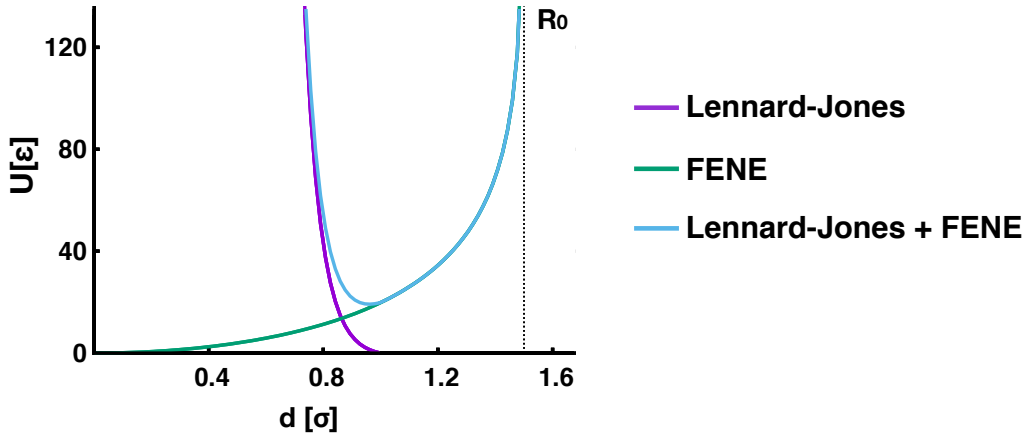


Fig. 2.2: Plot of Lennard-Jones potential (purple) and FENE potential (green) as a function of the beads distance d. Their sum is shown in blue.

An additional energy potential term is often added to model the bending rigidity of real polymers, like for example DNA [49]. For the purpose of our study, this potential is not necessary.

2.2 Langevin dynamics

The dynamics of the flexible chain introduced above, is studied by integrating numerically the equations of motion described by the *Langevin dynamics* [50]

Specifically, the position $\boldsymbol{r}_i$ of the i^{th} bead evolves accordingly with the equation:

$$m\ddot{\boldsymbol{r}}_i = -\gamma\dot{\boldsymbol{r}}_i - \boldsymbol{\nabla}_i U_i + \boldsymbol{\eta}_i \tag{2.3}$$

where γ is the friction coefficient, U_i is the potential acting on the i^{th} particle, $\boldsymbol{\eta}_i$ is a stochastic force acting on each bead and m is the mass (identical for all beads).
The stochastic noise term $\boldsymbol{\eta}_i$ is assumed to be isotropic, non-correlated in the dynamics time scale, and Gaussian distributed:

$$\begin{cases} \langle \eta_{i,\alpha}(t) \rangle = 0 \\ \langle \eta_{i,\alpha}(t)\, \eta_{j,\beta}(t') \rangle = 2K_B T\gamma\delta_{ij}\delta_{\alpha\beta}\delta(t-t') \end{cases} \tag{2.4}$$

where α, β and i, j pairs of indices respectively represent the Cartesian coordinates and the particle indices; δ_{ij} is the Kronecher delta and $\delta(t-t')$ the Dirac delta.
The eq. 2.3 describes the Brownian motion of particles in a medium, that is not otherwise explicitly modelled. Specifically, the terms $-\gamma\dot{\boldsymbol{r}}_i$ and $\boldsymbol{\eta}_i$ represent the dissipative and the stochastic force due to the collisions with the medium. The continuous injection and dissipation of energy due to these two terms, results in a dynamics that asymptotically returns the canonical (constant temperature) ensemble.
In case of absence of the external potential U_i, the mean displacement of the chain's centre of mass (CoM) is zero and the motion is characterised by the mean square displacement that, for time scales much larger than γ/m, has an asymptotic linear dependence on time:

$$\left\langle \Delta\boldsymbol{r}^2_{CoM}(t) \right\rangle_{t\to\infty} \simeq \frac{6K_B T}{\gamma N} t = 6D_0 t \tag{2.5}$$

where the diffusion coefficient D_0 goes as $1/N$, where N represents the length of the chain.
The Langevin dynamics is numerically integrated using LAMMPS simulation package [51].

2.3 Simulation units and time mapping

We will typically express our observables in *Lennard-Jones units*. In particular, distances will be given in units of the beads nominal diameter σ and energies in units of the thermal energy $\epsilon = K_B T$. The time scale will be expressed in *Lennard-Jones time*:

$$\tau_{LJ} = \sigma\sqrt{\frac{m}{\epsilon}} \tag{2.6}$$

Following the choice of ref. [48], customarily adopted in coarse-grained models, the friction coefficient will be set equal to:

$$\gamma = \frac{\sqrt{m\epsilon}}{2\sigma} \tag{2.7}$$

while the integration time step is set to $\Delta t = 0.012\ \tau_{LJ}$.
The simulation time can be approximately mapped to real time units via the Stokes' equation that relates the friction coefficient γ and the medium viscosity μ for a spherical particle of diameter σ:

$$\gamma = 3\pi\mu\sigma \tag{2.8}$$

By equating the expressions in eq. 2.8 and 2.7 we obtain:

$$\sqrt{m} = \frac{6\pi\mu\sigma^2}{\sqrt{\epsilon}} \tag{2.9}$$

and by substituting into eq. 2.6, we get:

$$\tau_{LJ} = \frac{6\pi\mu\sigma^3}{\epsilon} \tag{2.10}$$

where $\mu = 10^3\ Kg/m$ in case of water, $T = 300\ K$ at room temperature, and $\epsilon = K_B T = 4 \cdot 10^{-21}\ J$.
Therefore, for example, for $\sigma = 1.0\ nm$ we obtain $\tau_{LJ} = 4.7\ ns$.

Chapter 3

Sorting ring polymers by knot type using modulated nanochannels

In this chapter, I will present a theoretical and computational study to explore how the diffusive properties of ring polymers inside spatially modulated nanochannels are affected by the complex interplay between channel parameters, polymer length and polymer topology.
The aim is to check whether spatially modulated channels could, in principle, be used to sort ring polymers by knot type.
The content of this chapter is mainly based on the publication of ref. [21].

3.1 Background and motivation

In recent years several studies have characterised how polymer properties are affected by the inevitable presence of knots [5, 52]. They include both experimental and computational investigations focused on various properties like tensile strength [53], elastic response [54–56] and pore translocation capabilities [57–59].
A fundamental step for understanding the effects of knots on polymers' mechanical properties is studying the incidence of knots [60] and this, in turn, depends on the capability of establishing the knotted state of a given molecule [61, 62]. One possible challenging way to do it is by sorting the molecules by their knot type.
Presently, the most used technique for sorting knotted molecules by topology is gel electrophoresis [18, 19, 63–68]. This technique, developed around the 1930s, is typically used to sort biological macromolecules by their size, but it has also been used for the topological sorting of small viral DNA molecules since the 1970s [69, 70]. In this technique, DNA molecules, that are negatively charged, are driven by an electric field through the mesh of an agarose gel. The electrophoretic mobility of these molecules depends on their capability to pass through the gel obstacles and this ultimately depends on their intra-molecular entanglement.
The technique was first used by Wang and co-workers for analysing the knotted state of the $P4$ phage DNA [71]. $P4$ is a bacteriophage with a 10 kb genome, and it is usually studied both in its *wild type* and in the *tailless mutant* forms.

The latter lacks the characteristic tail that in the wild type is used to deliver the genome during the bacterial infection. Both wild-type and tailless genomes have 19-nucleotides long cohesive ends [72, 73] that lead to the formation of circular molecules. However, in wild types, the two cohesive ends of the DNA are not bonded because one end of the genome remains rooted close to the tail and therefore it can circularise only when in solution. On the contrary, in tailless mutants, the entire genome is reeled inside the capsids so that circular DNA molecules could result also from the stochastic encounter and subsequent annealing of the two ends inside the capsid.
Using gel electrophoresis, Roca and co-workers were able to prove that the incidence of knots is very high in tailless mutant DNA ($> 95\%$) compared to mature phage DNA (47%) [74] and to randomly cyclised DNA in solution (3%) [14]. Furthermore, they also characterised the knot types in the genome (see Fig. 3.1a), revealing that over 97% of the knots in tailless mutants were indeed very complex (being characterised by a crossing number greater than 10) [20]

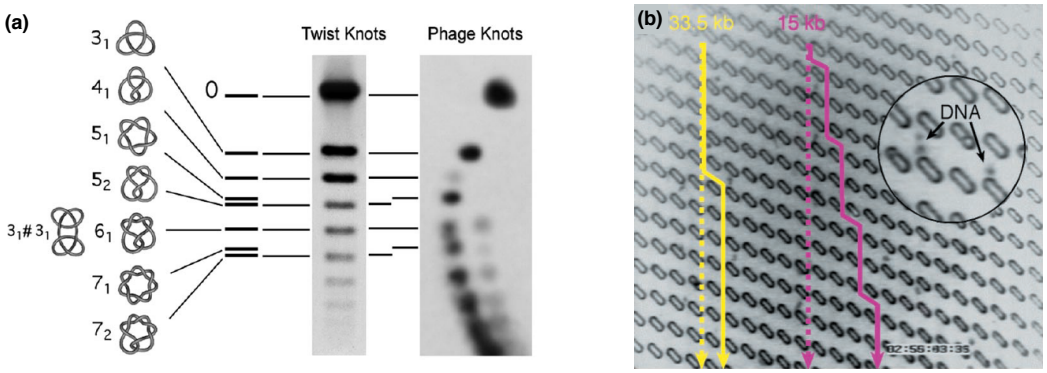


Fig. 3.1: **a)** Sorting of circular DNA molecules of tailless *P4* phage by knot type with two-dimensional gel electrophoresis (see ref. [20]). **b)** Sorting of linear DNA molecules by length using their diffusion properties through nano-arrays (see ref. [75]).

Gel electrophoresis, however, is limited regarding the range of DNA lengths addressable for the topological sorting. First, it can be used only to profile the knot type of relatively short circular molecules. In fact, if the molecule is longer than 10 *kb*, the strain developed at the gel obstacles is large enough to break the labile non-covalent circularisation of the chain. Secondly, gel electrophoresis resolution also degrades with chain length, especially for knots with a crossing number larger than 10 [20]. How to overcome these limitations is currently an open problem.
Over the years there have been advancements in microfabrication and micromanipulation techniques, that led to the development of new methods for sorting molecules by size or shape. Differently from gel electrophoresis, polymer chains are driven through arrays of posts or cavities by fluid or electric fields [76–80]. It was shown that the suitable interplay between intrinsic and extrinsic length-scales results in a chain mobility that depends on the polymer contour length, see Fig. 3.1b, or on its linear *versus* circular character [75, 81–85].
In particular, these studies included also the use of nanoslits or modulated

nanochannels as devices for the sorting [76, 84, 86, 87]. For the case of modulated nanochannels, the sorting does not rely on the pinning or jamming of the polymer against localised defects, but rather on entropic effects arising from the interplay of the intrinsic polymer size (e.g. its radius of gyration) and the size of the cavities. These entropic effects are advantageous compared to the pinning against localised defects because they are more easily adapted to longer polymer chains.
The possible extension to use modulated nanochannels or slits as setups to sort molecules by topology has been recently suggested by Micheletti and Orlandini. They used Monte Carlo simulations with DNA coarse-grained models to study equilibrium properties of knotted polymers confined in channels and slits [4, 88–90] and they found that polymers with different knotted states show different metric properties under spatial confinement. This result suggests that modulated nanochannels could, indeed, be used to confer a different mobility to molecules with different metric properties, and hence knot types.
Motivated by these results, here [21] we use molecular dynamics simulations and theoretical approaches to explore the possibility to perform the sorting of fully flexible polymers by their knot type using spatially and periodically modulated nanochannels.
As shown in the next section, a minimalistic chain-of-beads model is considered and the chains properties in a free diffusion dynamics are explored. Further refinements of the model are included in the last sections (active driving force) and pointed as possible future perspectives (hydrodynamics effects).

3.2 Model and methods

3.2.1 The model

As a model system we consider fully-flexible chain of beads confined inside a periodically modulated channel, as shown in Fig. 3.2.

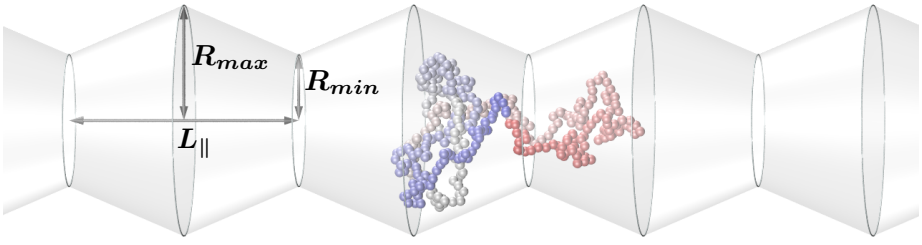


Fig. 3.2: Snapshot af a 4_1-knotted ring of $N = 350$ beads inside the modulated channel (geometry: $R_{min} = 8\sigma$, $R_{max} = 14\sigma$ and $L_{\parallel} = 2R_{max}$). Adapted from [21].

The modulated channel consists of the repeated succession of the same modular chamber. One such chamber is generated by the mirror juxtaposition of truncated cones. Hence, the chamber size is defined by three length scales: the minimum and maximum radii of its transverse (circular) sections, R_{min} and

R_{max} and the longitudinal span, $L_{\parallel}$.
The ring polymer is modelled as a fully-flexible closed chain of N beads as introduced in Chapter 2. The potential energy acting on the i^{th} bead of the chain consists of four additive terms:

$$U_i = U_{FENE}\left(d_{i,i-1}\right) + U_{FENE}\left(d_{i,i+1}\right) + \sum_{j=1}^{N} U_{LJ}\left(d_{i,j}\right) + U_{LJ}\left(D_i\right) \tag{3.1}$$

where U_{FENE} and U_{LJ} represents respectively the chain's connectivity (between consecutive beads) and the steric interactions introduced in Chapter 2. In eq. 3.1, $d_{i,j}$ denotes the distance between beads i and j, and D_i the distance between the bead i and the wall of the channel. Potentials parameters are set to their default values, as described in Chapter 2.

3.2.2 Molecular dynamics simulations

The dynamics of the system evolves according to a constant temperature Langevin dynamics with no hydrodynamic interactions. The equation parameters and the integration time step were set to their customary values, see Chapter 2.
Here, we consider rings made of $100 \leq N \leq 800$ beads and with three different topologies, unknotted (0_1), as well as tied in the trefoil (3_1) and figure-of-eight (4_1) knots. Simulations are started from a fully elongated ring configuration aligned on the channel axis (and spanning more than one chamber), which is next equilibrated for $\sim 10^6\ \tau_{LJ}$. Simulations time is $\sim 10^7\ \tau_{LJ}$.

3.2.3 Observables

The main observable monitored during the simulation is the position of the chain centre of mass (CoM) projected along the channel axis, x, also referred to as the *longitudinal position*:

$$r_{CoM_{\parallel}} = \frac{1}{N}\sum_{i=1}^{N} r_{i_{\parallel}} \tag{3.2}$$

To quantify the size of the chain, we use the *gyration radius*. It is defined as the root mean square distance of the beads from the chain's centre of mass:

$$\langle R_g \rangle = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left\langle \left(\boldsymbol{r}_i - \boldsymbol{r}_{CoM}\right)^2 \right\rangle_t} \tag{3.3}$$

where $\langle ... \rangle_t$ refers to the average over the trajectory.
The gyration radius in bulk for unknotted chains, as well as tied in trefoil and figure of eight knots, in the range $100 \leq N \leq 300$ is shown in Fig. 3.3.
A further quantity that we use is the chain extension in the longitudinal direction, defined as the *average longitudinal span*:

$$\left\langle span_{\parallel} \right\rangle = \left\langle \max_{i\in[1,N]}\left(r_{i_{\parallel}}\right) - \min_{j\in[1,N]}\left(r_{j_{\parallel}}\right) \right\rangle_t \tag{3.4}$$

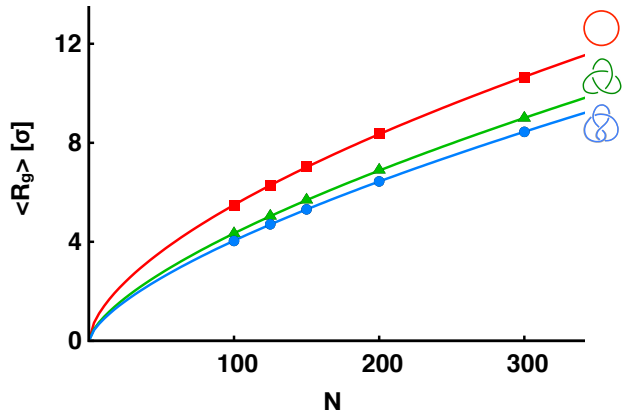


Fig. 3.3: Average gyration radius of unconstrained rings with contour lengths $100 \leq N \leq 300$ and three different topologies (squares: 0_1, triangles: 3_1 and circles: 4_1). The solid lines are power law fits of the type aN^b. Statistical errors are estimated from the block analysis of a single simulation of length $6 \cdot 10^6$ τ_{LJ} and are smaller than the representative symbol.

In addition to these three metric observables we also measured the contour length of the knotted portion, as well as its transverse size (maximum distance of any two chain beads) measured perpendicularly to the channel axis. The latter will also be referred to as *transverse knot span*.
The knotted region was identified by searching for the smallest portion of the chain that has the same topology of the entire chain once closed with the minimally-interfering closure scheme, see Chapter 1 [33].

3.2.4 Setting channel parameters

The geometry of the channel $R_{max} = L_{\parallel}/2 = 14\sigma$, was chosen to be comparable to the mean radius of gyration of unconstrained knotted and unknotted rings at the intermediate length, $N = 300$, as shown in Fig. 3.3.
As we will see later, the pore size, $R_{min} = 8\sigma$, was chosen as the best size that allows to discriminate the diffusion coefficient of unknotted and knotted chains of $N = 100$ beads (Fig. 3.16).

3.3 Diffusion in spatially modulated channel

As a first step we consider the time evolution of the chain's centre of mass position projected on the channel axis, x. A typical trajectory for a chain is shown in Fig. 3.4b. Two dynamical regimes are clearly visible: the motion is mostly confined within a single chamber for time scales much smaller than 10^6 τ_{LJ} while jumps, or transitions, between neighbouring chambers are commonly observed over longer time scales.
In order to have a better insight, we calculate the longitudinal mean square displacement of the centre of mass as a function of the time lag τ:

$$MSD_{\parallel}(\tau) = \left\langle \left(r_{CoM_{\parallel}}(t+\tau) - r_{CoM_{\parallel}}(t) \right)^2 \right\rangle_t \tag{3.5}$$

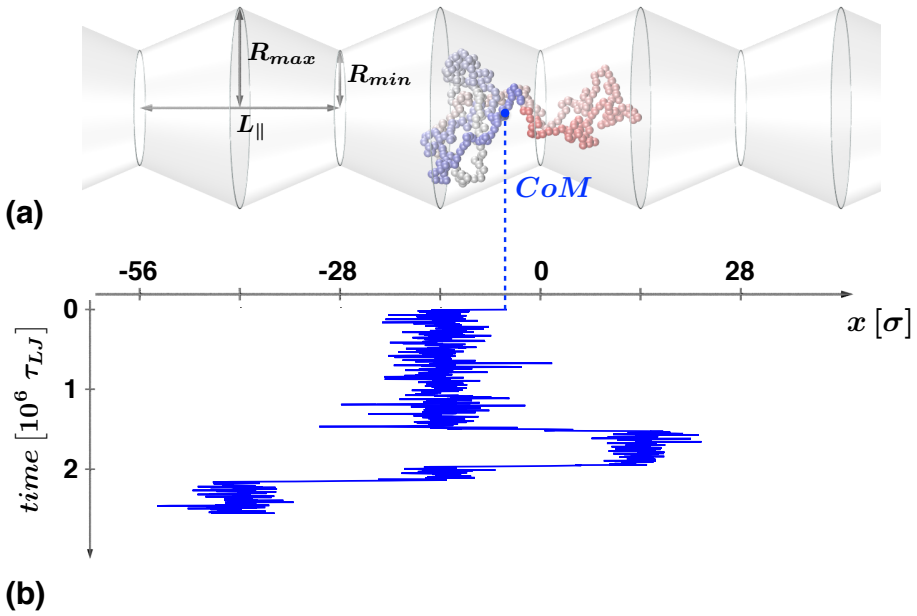


Fig. 3.4: Snapshot af a 4_1-knotted ring of $N = 350$ beads inside the modulated channel (geometry: $R_{min} = 8\sigma$, $R_{max} = 14\sigma$ and $L_{\parallel} = 2R_{max}$). The ring's centre of mass is highlighted and a typical time evolution of its position on the longitudinal axis of the channel, x, is shown. Taken from [21].

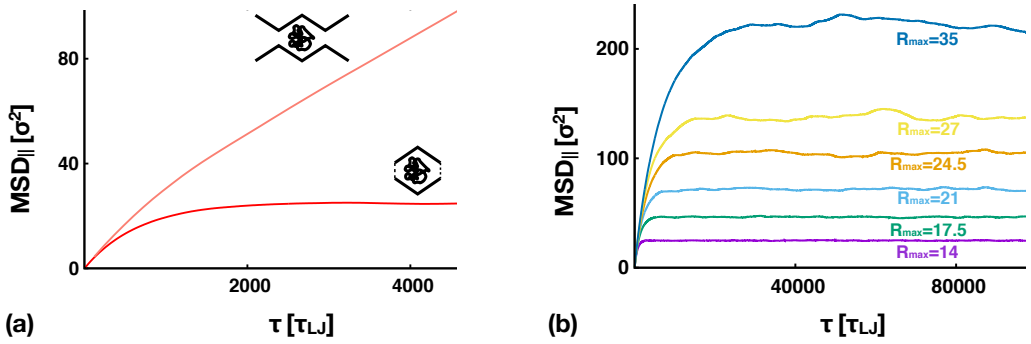


Fig. 3.5: **(a)** Longitudinal mean square displacement ($MSD_{\parallel}$) as a function of the time lag τ for an unknotted chain of $N = 100$ beads. The light-red curve represents the MSD of the ring in the periodically-modulated channel. The red curve instead represents the MSD for the chain trapped inside a single isolated (i.e. walled) chamber of the modulated channel. The same quantity is shown in panel **(b)** for walled chambers with different values of R_{max}. Taken from [21].

The two regimes are illustrated in Fig. 3.5a, which shows the time dependence of the mean square displacement of the centre of mass for a ring with $N = 100$ beads inside a single walled chamber (the chain is not allowed to translocate into adjacent chambers) and inside the "infinitely" extended modulated channel.

In the first case, the MSD plateaus at time lags larger than 2000 τ_{LJ}, which therefore identifies the characteristic intra-chamber exploration time for the ring and hence the onset of the inter-chamber diffusive regime. The chamber

exploration time clearly increases with the chamber size, as shown in Fig. 3.5b. In the latter case the asymptotic longitudinal motion is well-consistent with normal diffusion, as expected, since the jumps in adjacent chambers are stochastic.

3.4 Diffusion of unknotted and knotted rings

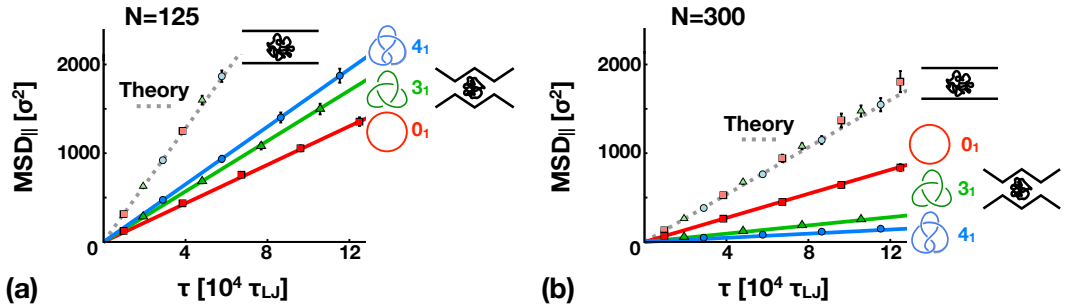


Fig. 3.6: Mean square longitudinal displacement $MSD_{\parallel}$ versus time lag τ for chains of **(a)** 125 beads and **(b)** 300 beads and different topologies: 0_1, 3_1 and 4_1. Averages are taken over 10 independent trajectories. The solid lines are direct proportionality best fits for modulated channel data (saturated colours). The dashed line is the theoretical dependence for the uniform channel case (light colours). Taken from [21].

In order to check if modulated channels can indeed be used to sort equally-long molecules by knot type, we simulated chains of different topologies and lengths both in uniform ($R_{min} = R_{max} = 14\sigma$) and modulated channels for a time of around 10^7 τ_{LJ} . In particular we considered unknotted chains, as well as tied in a trefoil (3_1) and figure-of-eight (4_1) knots with a chain length of $N = 150$ and $N = 300$ beads. Their $MSD_{\parallel}$ is shown in Fig. 3.6. In case of diffusion inside the uniform channel, the motion is insensitive to the knot type both for shorts and long chains, as expected in absence of hydrodynamics effects. In fact, the uniform channel provides a confining force perpendicular to the longitudinal channel axis, which is the direction of interest for the motion. It is therefore no surprising that the mean square displacement follows the diffusive trend of eq. 2.5 ($D_0 = K_B T/\gamma N$).

A qualitatively different behaviour is instead found in the modulated channel case. In fact, for both $N = 125$ and $N = 300$ the chain mobility varies significantly with knot type. The most surprising feature is that the mobility ranking of knots changes dramatically with chain length. In fact, at $N = 125$, the fastest diffuser is the most complex considered knot (4_1), while at $N = 300$ it is the unknot.

To clarify the physical mechanisms that cause this dramatic inversion of the mobility ranking, we analysed the diffusion coefficient D of rings with 0_1 and 4_1 topologies varying N in the 100 to 800 range. Again, we performed simulations both in uniform and modulated channels.

Results, given in Fig. 3.7, show that the diffusion coefficients in the modulated channel for both topologies depart significantly from the $1/N$ trend of the uniform channel case, which provides an upper bound for D.

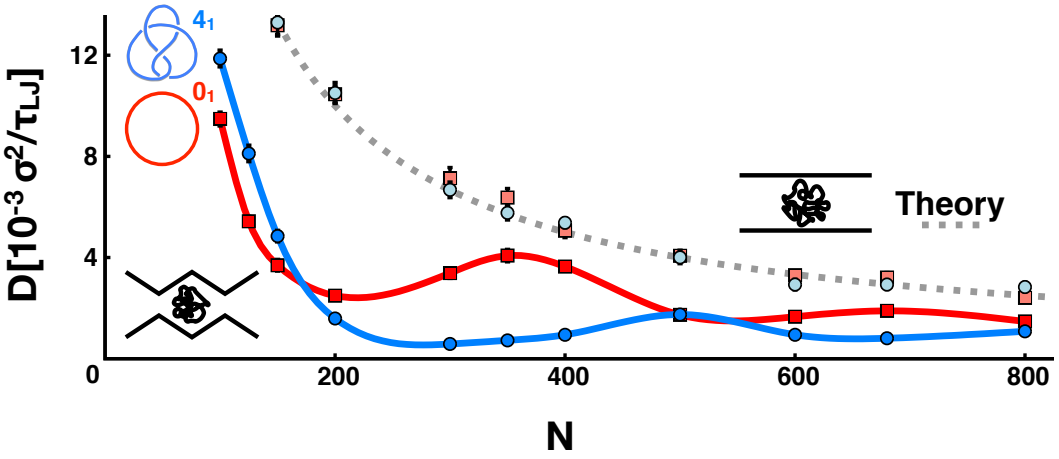


Fig. 3.7: Longitudinal average diffusion coefficient, D, for unknotted and 4_1-knotted rings of 100 to 800 beads in modulated and uniform channels (in saturated and light colours, respectively). The solid curves (splines) are provided as visual guides. The dashed line is the theoretical expectation for uniform channels, $D_0 = 2\sigma^2/N\tau_{LJ}$. Averages are taken over 10 independent trajectories. Taken from [21].

In particular, D presents a non-monotonic oscillatory behaviour as a function of the chain's length and appreciable differences are observed between unknotted and knotted chains. Firstly, the minima and maxima of the oscillations are not in register for the two topologies. The mismatch of minima at the shortest chain lengths causes the D curves to cross at $N \sim 150$. This, in turn, produces the inversion of the knot-dependent mobility ranking noted in Fig. 3.6. Secondly, the diffusion coefficient of knotted rings is much suppressed for $N > 200$, differently from the unknotted counterparts.
In the following, we will first present a qualitative analysis to identify the key features related to oscillations and we will explore their dependence on topology. Then, we will perform a more quantitative analysis based on the Fick-Jacobs approximation.

3.5 The oscillation of D in unknotted chains

As we discussed above, chains diffuse inside a single chamber and, passing through the pore, are stochastically able to translocate to an adjacent chamber. If we restrict on considering the longitudinal position of the CoM, we can represent the "obstructing" effect of the pore with a *free energy* curve, as sketched in Fig. 3.8. This is the free energy barrier that the CoM needs to overcome to jump to a neighbouring chamber. As we will see in this section, the free energy is the key quantity to understand the oscillations of the diffusion coefficient D. A part for an additive term, which amounts to an overall free energy shift, the free energy can be expressed as a function of the probability $P(x)$ of finding the CoM at longitudinal position x as:

$$F(x) = -K_B T \log\left[P\left(x\right)\right] \tag{3.6}$$

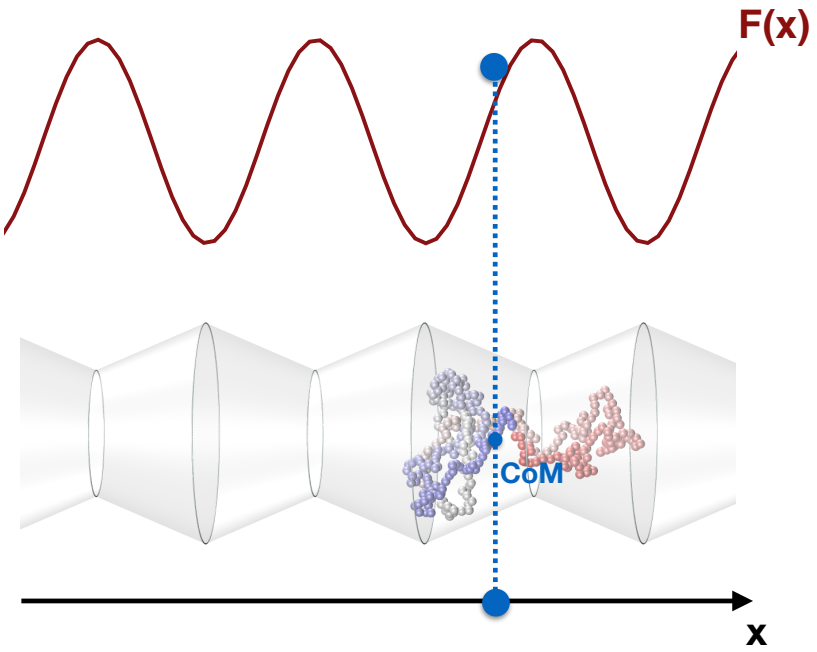


Fig. 3.8: Sketch of the free energy barrier that the CoM needs to overcome to pass through the pores.

The probability $P(x)$ is obtained through the histogram of the longitudinal position, with x mapped periodically in the $[0, L_{\parallel}]$ interval.

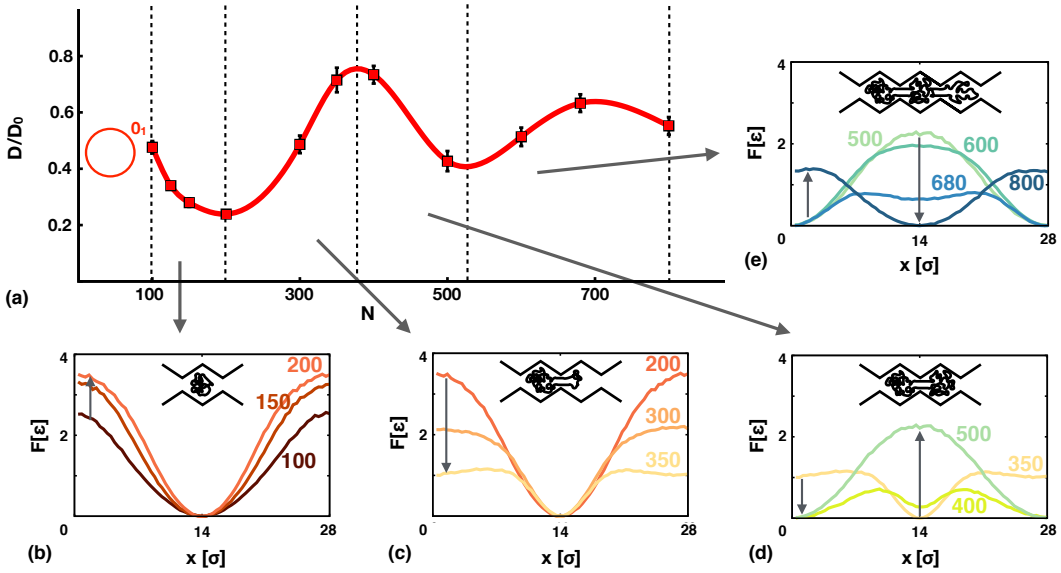


Fig. 3.9: **a)** Normalised longitudinal diffusion coefficient, D/D_0, for unknotted rings of 100 to 800 beads as computed directly from simulated trajectories. The curves (splines) are provided as visual guides. Averages are taken over 10 independent trajectories. **b-e)** Free energy profiles $F(x)$ for different intervals of N. Distributions are calculated from uncorrelated configurations of 10 independent trajectories. Adapted from [21].

Here, we consider the normalised diffusion coefficient D/D_0 for unknotted rings as a function of the chain length alongside the relative variation in the free energy barriers, as shown in Fig. 3.9. The normalisation factor $D_0 = K_B T/\gamma N$ is the diffusion coefficient in the uniform channel.
The pronounced oscillatory profile illustrates more vividly than Fig. 3.7 the deviations from the uniform channel case, where D/D_0 would be constant and equal to 1. Here, instead, the adimensional ratio D/D_0 for unknotted rings takes on values that range from 0.2 to 0.7 and presents distinct regimes, as shown in Fig. 3.9a.

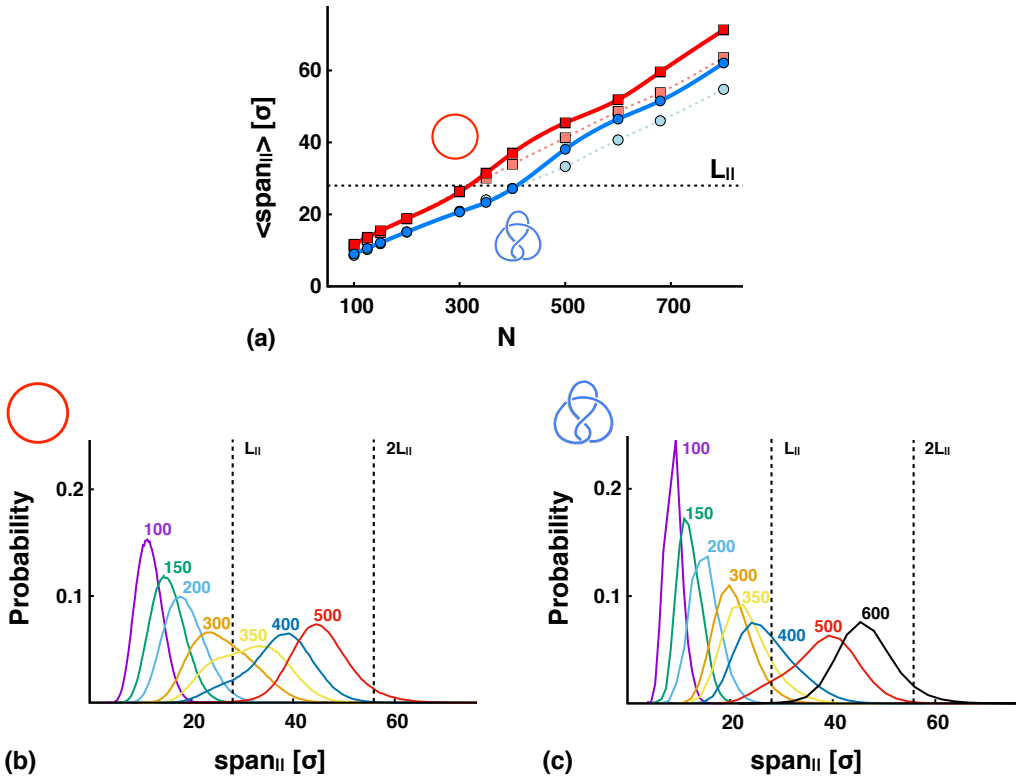


Fig. 3.10: **(a)** Saturated colours indicate the average longitudinal span of unknotted (red squares) and 4_1-knotted rings (blue circles) in modulated nanochannels as a function of N. Data for uniform channels are shown in lighter colours. The saturated and dashed-light (spline) curves are provided as visual guides. The black dashed line marks the chamber longitudinal span, $L_\parallel$. Averages are taken over uncorrelated frames of 10 independent trajectories. **(b-c)** The probability distributions of the longitudinal span of unknotted **(b)** and 4_1-knotted **(c)** rings for different values of N. The dashed lines mark the values for $L_\parallel$ and $2L_\parallel$. Distributions are calculated from uncorrelated configurations of 10 independent trajectories. Adapted from [21].

One observes that for chains with a gyration radius much smaller than the chamber, $100 \leq N \leq 200$, the centre of mass is mostly located at the centre of the chamber, where F is at its minimum, while the free-energy barriers at the edges increase with N, see Fig. 3.9b. The latter property is consistent with the expected lower rate at which longer chains can move across adjacent chambers. This regime holds up to $N \sim 200$ when chains become long enough to protrude into a neighbouring chamber because their average longitudinal span exceeds $L_\parallel$, as shown in Figs. 3.10a.

Further increments in N increase the fraction of rings that straddle two chambers, see Fig. 3.10b and 3.11a. The cross-chamber barrier is accordingly progressively reduced, see Fig. 3.9c, and this facilitates the trans displacement. The resulting increase of D/D_0 holds up to $N \sim 350 - 400$, when the ring portions, or blobs, in neighbouring chambers are typically of equal size and when the probability of the ring to occupy a single chamber becomes smaller than occupying two chambers, see Fig. 3.11a. The peak at $N \sim 350 - 400$ corresponds to the chain length at which a free energy barrier starts to develop in the middle of the chamber, therefore further suppressing the chances of finding the centre of mass at that position, see Fig. 3.9d. In this situation the diffusive motion becomes hindered by the balanced size of the two "blobs" that compete for growing and therefore also for pulling the chain across the chambers.

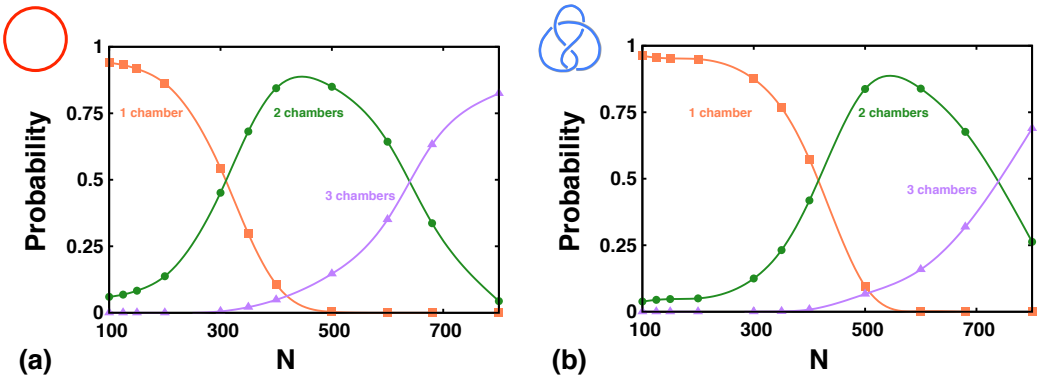


Fig. 3.11: Probability distribution that a ring occupies 1 chamber (orange curve), 2 chambers (green curve) or 3 chambers (violet curve). The data probabilities are shown as a function of N for **(a)** unknotted rings, as well as **(b)** tied in a figure-of-eight knots. The number of occupied chambers is calculated conservatively by neglecting cases where the longitudinal protrusions into the chambers are less than $L_{\parallel}/10$. Probabilities are calculated from uncorrelated configurations of 10 independent trajectories. Taken from [21].

This modulation pattern associated to the chain accommodating from one to two chambers is repeated, even if with reduced magnitude, for longer chains expanding from two to three chambers and so on, see Fig. 3.9e.
We can therefore conclude that the D *versus* N oscillations in Fig. 3.7 and 3.9 are related to the interplay between the chain and the chamber size, effectively described by the confining free energy profiles. Knotted rings show a similar oscillatory behaviour in D, as we can see from Fig. 3.7 and from their longitudinal span $L_{\parallel}$ or the number of occupied chambers shown respectively in Figs. 3.10, and 3.11b . However, differences in the diffusion of knotted rings emerge due to the presence of knots, as we discuss below.

3.6 The effects of knots

The two major differences observed in the diffusion of knotted and unknotted rings are: a mismatch in the oscillating trends of the diffusion coefficient and a significant reduction of the same for knotted rings with $N > 200$ (see Fig. 3.7). The shift between unknotted and knotted rings diffusion coefficients is found to

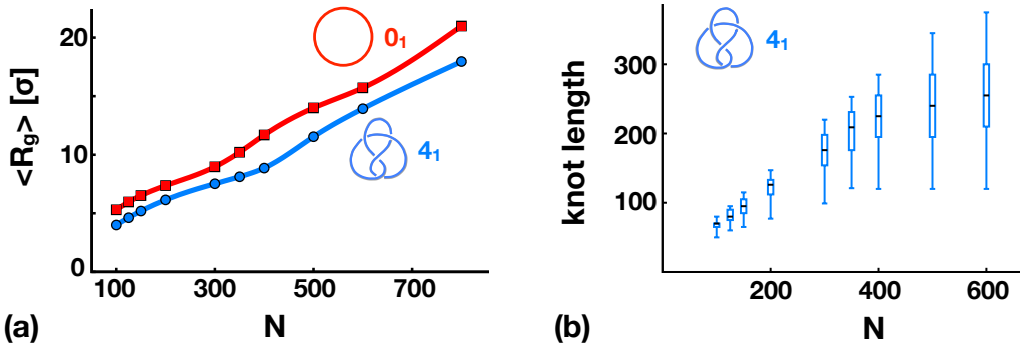


Fig. 3.12: **(a)** Average radius of gyration of rings of 100 to 800 beads with unknotted and 4_1 topologies inside the modulated channel. The curves (splines) are provided as visual guides. Averages are taken over uncorrelated configurations of 10 independent trajectories. **(b)** Length-dependence of the knot length for 4_1-knotted rings inside the modulated channel. The black line marks the median value, the box is delimited by the 25^{th} and 75^{th} percentiles, while bars mark the 2^{nd} and the 98^{th} percentiles. Distributions are calculated from uncorrelated configurations of 10 independent trajectories. Adapted from [21].

be about 100 beads, see Fig. 3.7. Interestingly, one observes that the same difference is found between the gyration radius of unknotted and knotted chains: the gyration radius of knotted rings is about equivalent to that of unknotted ones shortened by about 100 beads, see Fig. 3.12a. This suggests that the mismatch of D is related to the effective contour length of knotted rings, which is decreased with respect to the unknotted ones, because part of the chain is used up to maintain its non-trivial topological state. The idea is indeed confirmed by checking that this length reduction is also of the same order as the most probable knot length of the confined chains for the considered range of N, see Fig. 3.12b.

For knotted chains with more than 200 beads we then see a significant reduction of the diffusion coefficient.

Fig. 3.13 shows that for $N > 200$ the average transverse knot span, which increases with N, becomes larger than the diameter of the connecting pore, suggesting that knots can directly hinder the translocation process when they are large enough. Ring migration across chambers is allowed only when atypical fluctuations of the transverse knot size make the latter smaller than the pore size. These fluctuations are represented by the tails of the distributions in Fig. 3.13b.

As a consequence, the knot hindrance causes a significant increase of the dwelling time of knotted rings with respect to the unknotted ones for sufficiently long chains, see Fig. 3.14.

The above discussion clarifies the diffusive motion in modulated channels. The complex dependence of D on chain length and topology, results from an interplay not only of the size of the chain and the size of the chamber, but also of the knot as well. In the next section, we will rationalise this result with a theoretical framework that describes the diffusion processes in confined geometries.

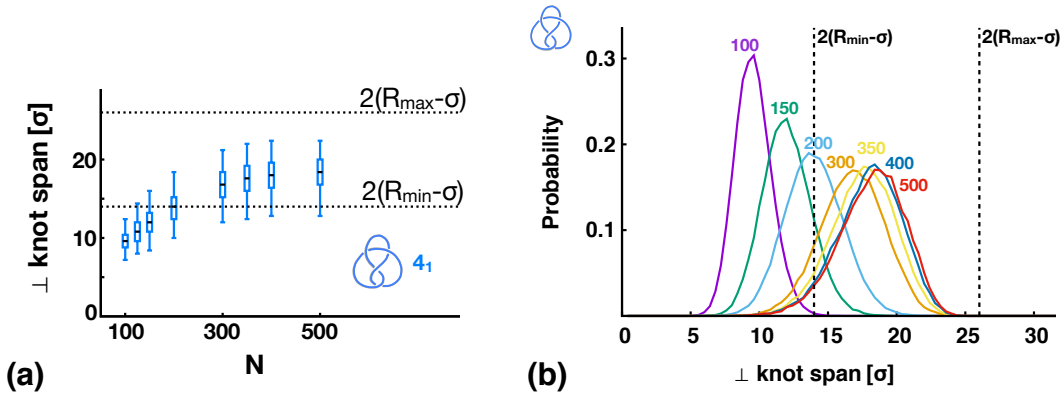


Fig. 3.13: **(a)** Length-dependence of the average transverse knot span for 4_1-knotted rings inside the modulated channel. The black line marks the median value, the box is delimited by the 25^{th} and 75^{th} percentiles, while bars mark the 2^{nd} and the 98^{th} percentiles. **(b)** Probability distribution of the transverse knot span for 4_1-knotted rings with $100 < N < 500$. The two dashed lines indicate the effective maximum and minimum transverse diameters of the modulate chamber. The maximum and minimum effective transverse sizes of the chamber are shown for reference with dotted lines. Distributions are calculated from uncorrelated configurations of 10 independent trajectories. Adapted from [21].

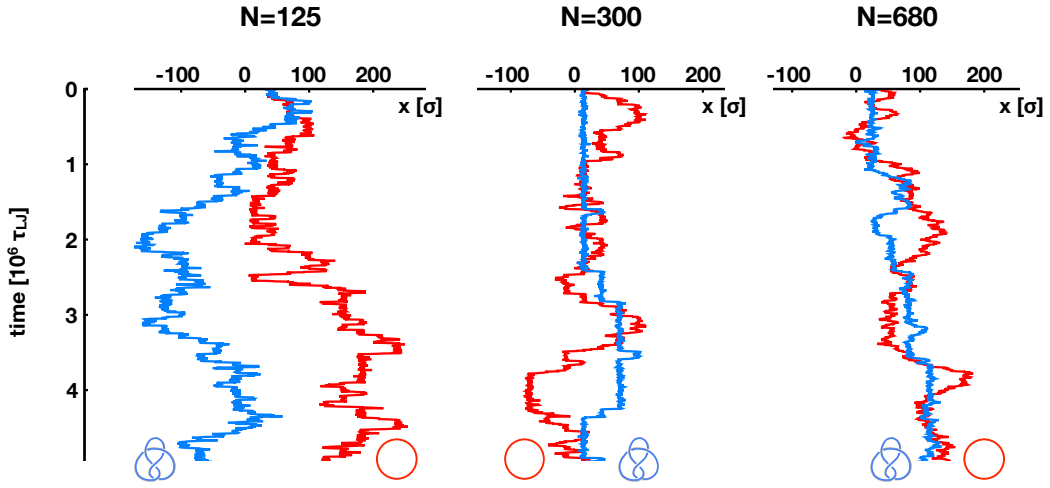


Fig. 3.14: Trajectories of the longitudinal CoM displacement of rings with $N = 125$, 300 and 680 beads and two different topologies (unknotted in red and figure-of-eight in blue). The hindrance to pore translocation introduced by knots for $N = 300$ and 680 is reflected in the longer intra-chamber dwelling times compared to unknotted rings with same lengths. Taken from [21].

3.7 The Fick-Jacobs approximation

We now interpret the oscillatory behaviour of D with the aid of the Fick-Jacobs approximation, a theoretical framework that quantitatively relates the diffusion coefficient to the free energy barrier generated by the confined geometry. We briefly overview the theory of this approach and then discuss its application to our system.

3.7.1 Fick-Jacobs theory

The *Fick-Jacobs approximation* or *theory* (FJ), firstly introduced by Jacobs in 1967 [91] and then re-organised by Zwanzig in 1992 [92], is a theoretical approach that simplifies the treatment of the quasi one-dimensional diffusive motion of single particles inside confined geometries. Here, in particular, we consider confinement due to periodic modulated nanochannels of circular section and period $L_{\parallel}$.
The FJ theory permits to coarsen the description by reducing the dimensionality of the system, keeping only the main direction of diffusion, namely the direction of the longitudinal axis of the channel, x. However, the dimensionality reduction is based on the assumption that the particle distribution in the transverse direction (y and z) equilibrates much faster than that in the unconstrained direction of diffusion (x).
Consider the Langevin equation for a single particle of mass m diffusing in a periodic modulated channel, where the confinement is described by the potential $U(\boldsymbol{r})$:

$$m\ddot{\boldsymbol{r}} = -\gamma\dot{\boldsymbol{r}} - \boldsymbol{\nabla}U(\boldsymbol{r}) + \boldsymbol{\eta} \tag{3.7}$$

where γ is the friction coefficient and $\boldsymbol{\eta}$ the stochastic force.
As known, the time evolution of the probability $p(\boldsymbol{r}, t)$ of finding the particle at position $\boldsymbol{r}$ is described by the *Smoluchowski equation*:

$$\partial_t p(\boldsymbol{r}, t) = \boldsymbol{\nabla}^2 \left[\left(D_0 + \frac{U(\boldsymbol{r})}{\gamma} \right) p(\boldsymbol{r}, t) \right] \tag{3.8}$$

that can be rewritten also in an equivalent form useful in the Fick-Jacobs theory:

$$\partial_t p(\boldsymbol{r}, t) = \boldsymbol{\nabla} \cdot \left[D_0 e^{-\beta U(\boldsymbol{r})} \boldsymbol{\nabla} \left(e^{\beta U(\boldsymbol{r})} p(\boldsymbol{r}, t) \right) \right] \tag{3.9}$$

Here, D_0 is the diffusion coefficient in bulk.
The *Fick-Jacobs equation* is derived from eq 3.9 by integrating out the transverse coordinates y and z and considering only the main direction of diffusion x. As mentioned before, the dimensionality reduction is based on the assumption of equilibration in the transverse direction.
The FJ equation therefore describes the time evolution of the probability along the longitudinal channel axis:

$$\partial_t p(x, t) = \partial_x \left[D_0 \; e^{-\beta A(x)} \partial_x \left(e^{\beta A(x)} p(x, t) \right) \right] \tag{3.10}$$

In eq. 3.10 the 3-dimensional potential $U(x, y, z)$ was replaced by the 1-dimensional confining potential $A(x)$, defined as:

$$e^{-\beta A(x)} = \int_{\mathbb{R}} dy \; dz \; e^{-\beta U(x,y,z)} \tag{3.11}$$

As mentioned above, the potential $U(x, y, z)$ describes the confinement due to

the channel and therefore is a box-like potential of the form:

$$U(x,y,z) = \begin{cases} 0 & \text{if } (x,y,z) \in \Omega(x) \\ \infty & \text{if } (x,y,z) \notin \Omega(x) \end{cases} \tag{3.12}$$

where $\Omega(x)$ is the allowed confined region for a given longitudinal position x. Hence, the definition of $A(x)$ becomes

$$e^{-\beta A(x)} = \int_{y,z\in\Omega(x)} dy\ dz\ = Q(x) \tag{3.13}$$

where $Q(x)$ represents the area of the transverse section of the channel at position x and since the channel has circular sections of radius $w(x)$, $Q(x) = \pi w^2(x)$. $Q(x)$ measures all the possible configurations available at a given position x, therefore $A(x)$ represents an entropic potential generated by the confining geometry:

$$A(x) = -\frac{1}{\beta}\log Q(x) \tag{3.14}$$

that, in absence of external forces, is periodic if the channel geometry is also periodic.
Note that the FJ theory holds also in transport motions where particles are subject to an external driving force f. In that case, the potential $A(x)$ has also an enthalpic term $-f\cdot x$.
A generalisation of the FJ eq. 3.10 was developed by Zwanzig [92], who introduced a position-dependent diffusion coefficient $D(x)$ to improve the approximation validity:

$$\partial_t p(x,t) = \partial_x\left[e^{-\beta A(x)} D(x)\partial_x\left(e^{\beta A(x)} p(x,t)\right)\right] \tag{3.15}$$

where $D(x)$ is given by

$$D(x) \simeq \frac{D_0}{1+\xi\ \left(\frac{dw(x)}{dx}\right)^2} \tag{3.16}$$

The coefficient ξ depends both on the confined geometry and the system dimensionality. In case of cylindrical symmetry in 3-dimension, $\xi = 1/2$.
The main quantity that we aim to calculate is the effective diffusion coefficient, that governs diffusion of long time scales (which is different than the local $D(x)$):

$$D_{eff} = \lim_{t\to\infty} \frac{\langle x^2(t)\rangle}{2t} \tag{3.17}$$

There are several ways to obtain the effective diffusion coefficient for motion in a 1-dimensional periodic potential. From the Fick-Jacobs equation, one can follow, for example, the procedure of Lifson and Jackson [93], which is based on the mean first passage time [94–96]. In either case, one obtains a simple

expression for the effective diffusion coefficient:

$$D_{eff} = \frac{1}{\langle e^{-\beta A(x)} \rangle_x \langle e^{\beta A(x)}/D(x) \rangle_x} = \frac{1}{\langle Q(x) \rangle_x \left\langle \frac{1}{Q(x)D(x)} \right\rangle_x} \tag{3.18}$$

where the average is calculated over one period $L_\parallel$ of the channel:

$$\left\langle e^{-\beta A(x)} \right\rangle_x = \int_0^{L_\parallel} \frac{dx}{L_\parallel} e^{-\beta A(x)} \qquad \left\langle \frac{1}{e^{\beta A(x)}/D(x))} \right\rangle_x = \int_0^{L_\parallel} \frac{dx}{L_\parallel} \frac{e^{\beta A(x)}}{D(x)} \tag{3.19}$$

3.7.2 Validity of Fick-Jacobs theory

As mentioned above, the reduction of dimensionality in the Fick-Jacobs equation relies on the assumption of equilibration in the transverse direction. An estimate of the conditions under which equilibration occurs can be made by analysing the two time scales involved in the problem.
One is the time scale τ_T necessary for the particle to diffusively cover the transversal section of the channel, and which depends on the position x:

$$\tau_T = \frac{w^2(x)}{2D_0} \tag{3.20}$$

The second time scale τ_L is the one that characterises the longitudinal diffusive motion and represents the time required to cover a period $L_\parallel$:

$$\tau_L \sim \frac{L_\parallel^2}{2D_0} \tag{3.21}$$

To have equilibration, the characteristic time scale associated to the transversal motion, must be much smaller than the one associated to the longitudinal motion. This, as a first approximation, can be achieved in relatively smooth channels:

$$\frac{\tau_T}{\tau_L} = \frac{w^2(x)}{L_\parallel^2} \sim \left(\frac{dw(x)}{dx} \right)^2 \ll 1 \tag{3.22}$$

3.7.3 Application to polymers diffusing in modulated channels

The application of Fick-Jacobs theory to our chain of beads is obtained by establishing the correspondence between the particle and the ring's centre of mass, and between the confining potential $A(x)$ and the free-energy $F(x)$ calculated from the probability of the centre of mass position (see eq. 3.6). This mapping can hold only approximately because, unlikely ideal particles, rings are spatially extended and hence the effective confining potential could vary on the ring scale. The effective diffusion coefficient of eq. 3.18 is given by

$$D_{eff} = \frac{D_0}{(1+\Gamma) \langle e^{-\beta F(x)} \rangle_x \langle e^{\beta F(x)} \rangle_x} \tag{3.23}$$

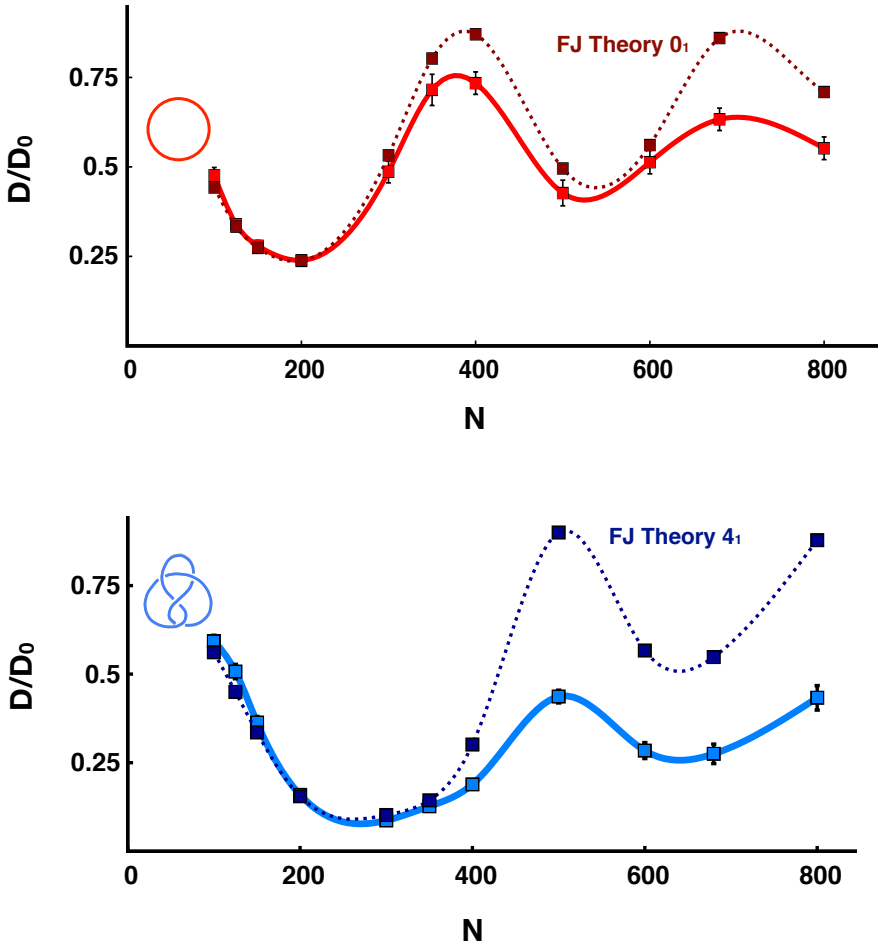


Fig. 3.15: Normalised longitudinal diffusion coefficient, D/D_0, for **(a)** unknotted rings and **(b)** 4_1-knotted rings of 100 to 800 beads as computed directly from simulated trajectories (bright red and blue respectively) and from the Fick-Jacobs approximation (dark red and dark blue respectively). The curves (splines) are provided as visual guides. Averages are taken over 10 independent trajectories. Adapted from [21].

where the local diffusion coefficient $D(x)$ of eq. 3.16 is obtained from the geometry of the modulated channel as

$$D(x) = \frac{D_0}{1 + \xi \left(\frac{dw(x)}{dx}\right)^2} = \frac{D_0}{1 + \frac{2}{L_\parallel^2}\left(R_{max} - R_{min}\right)^2} = \frac{D_0}{(1+\Gamma)} \tag{3.24}$$

Eq. 3.23 precisely contains the non-monotonic D *versus* N dependence that we aim to investigate. This is shown in Fig. 3.15, where the diffusion coefficient is compared with that computed from the Fick-Jacobs theory, both for unknotted and 4_1-knotted rings. The agreement of the curves holds well up to $N = 300$, when the average longitudinal span of the ring becomes comparable to $L_\parallel$, see Fig. 3.10. Beyond such chain length, the curves maintain the same qualitative trend but their quantitative agreement noticeably degrades. This is arguably because the rings become too extended to be represented by the sole centre

of mass in the FJ approximation. Furthermore, the quantitative agreement degrades more severely for knotted chains, because the obstructive role of the knot is not taken in account.
When chains are small, the FJ approximation also reproduces the effects of the pore on chains diffusion, as shown in Fig. 3.16. Here, D is shown as function of the pore size for chains of $N = 100$ beads. The agreement between theory and simulation is quantitatively good for both topologies.

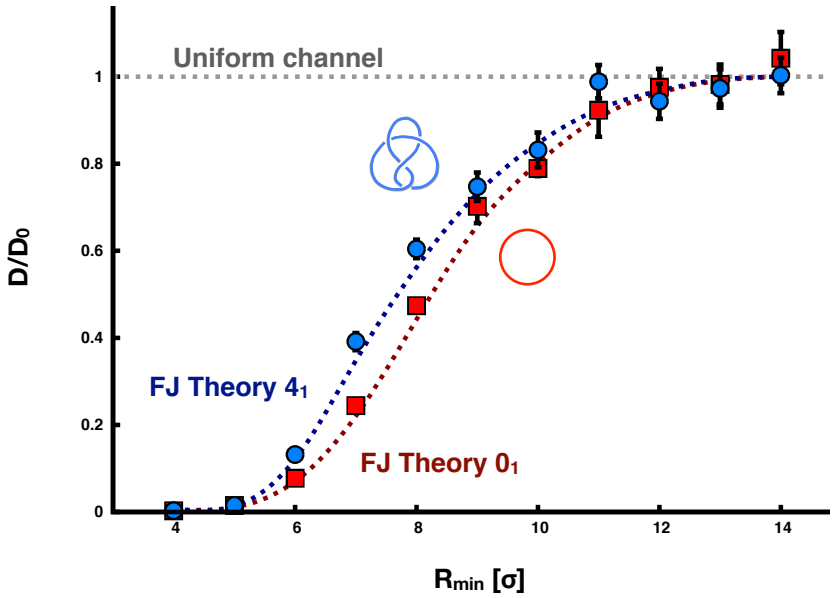


Fig. 3.16: Normalised diffusion coefficient D/D_0 as a function of the pore radius R_{min} for rings with $N = 100$ beads and two different topologies, unknotted (red squares) and figure-of-eight (blue circles) knotted rings. Dark-red and dark-blue dashed lines are the diffusion coefficients of the rings calculated using the Fick-Jacobs approximation. Averages are taken over 10 independent trajectories. Taken from [21].

We can therefore conclude that this approximation supports (with relative limitations) in a more quantitative manner the observations that we made in Section 3.5 about the relation between D and the confining free energy $F(x)$.

3.8 Optimal sorting condition for a given chain length

The above discussion clarifies that the complex dependence of D on chain length and topology, results from a rich interplay of the size of the chain, of the chamber, and of the knot as well. More importantly, this interplay can be used to guide the design of the modulated nanochannels to optimise their sorting capabilities for rings of given length.
To illustrate this point we note that with the chamber geometry considered so far ($R_{min} = 8\sigma$, $R_{max} = L_{\parallel}/2 = 14\sigma$) the largest dynamical range of D is obtained for $N \sim 300 - 350$, when the mobility of the unknotted rings is 6 times larger than 4_1-knotted ones, see Fig. 3.7.

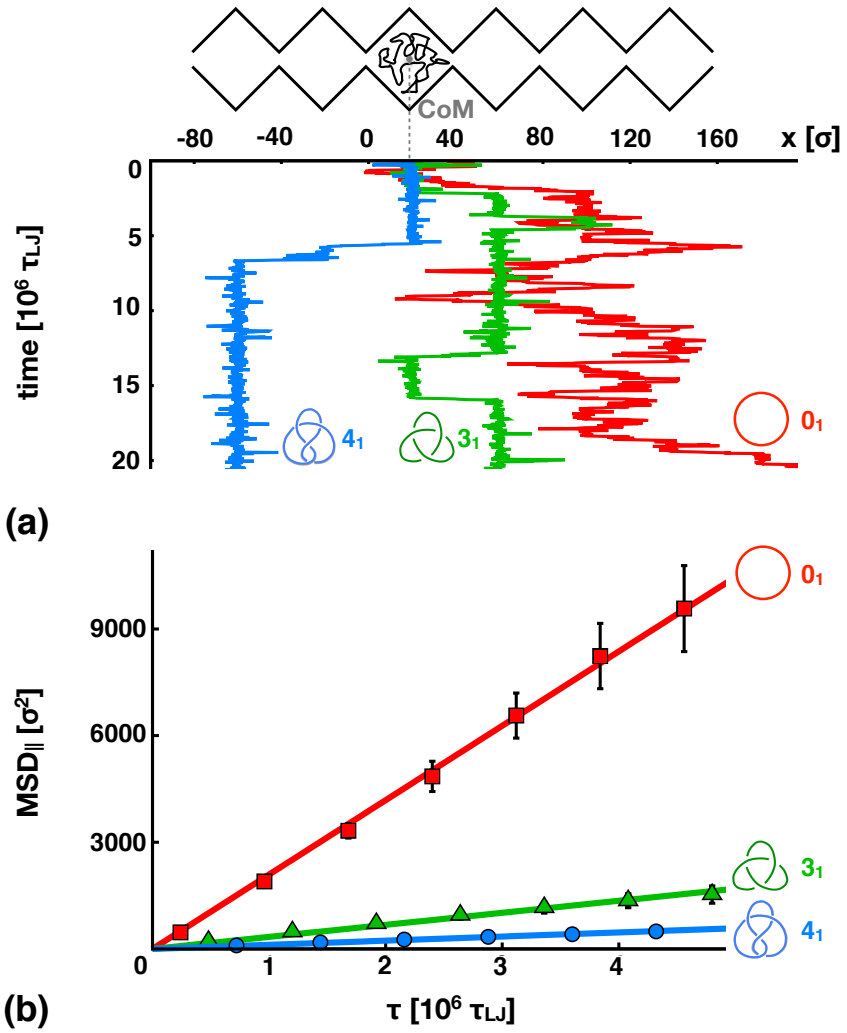


Fig. 3.17: **(a)** Typical trajectory for the longitudinal displacement of rings with 680 beads and different topologies inside the optimally-modulated nanochannel ($L_{\parallel} = 2R_{max} = 40\sigma$, $R_{min} = 8\sigma$). **(b)** Mean square longitudinal displacement of the rings versus time lag τ. Averages are taken over 10 independent trajectories. The solid lines are direct proportionality best fits. Adapted from [21].

For longer rings, of e.g. $N = 680$, the difference reduces to about a factor of 2. The sorting efficiency for these rings can still be dramatically improved upon by using a larger chamber, yielding a chain *vs* chamber span and volume ratio similar to the one found for $N \sim 300$. To meet these conditions for rings of $N = 680$ beads, we set $R_{max} = L_{\parallel}/2 = 20\sigma$, while we kept $R_{min} = 8\sigma$ because the N-dependence of the knot transverse size is mild for $N > 200$, see Fig. 3.12b.

Fig. 3.17 shows typical trajectories and evolutions of the mean square longitudinal displacement of these long chains inside the rescaled channel geometry. Results are shown both for rings knotted in trefoils as well as figure-of-eight knot. Unknotted rings now show a much higher diffusion coefficient than the knotted counterparts. In fact, for this optimised channel geometry, the diffusion coefficients of the 0_1 and 4_1 knots now differ by a factor of 20. This is a ten-fold

increase over the original case. Furthermore, different diffusion coefficients are observed also between the 3_1- and 4_1-knotted chains. This result opens up the possibility of sorting in respect of the topological complexity.

3.9 Sorting in presence of a driving force

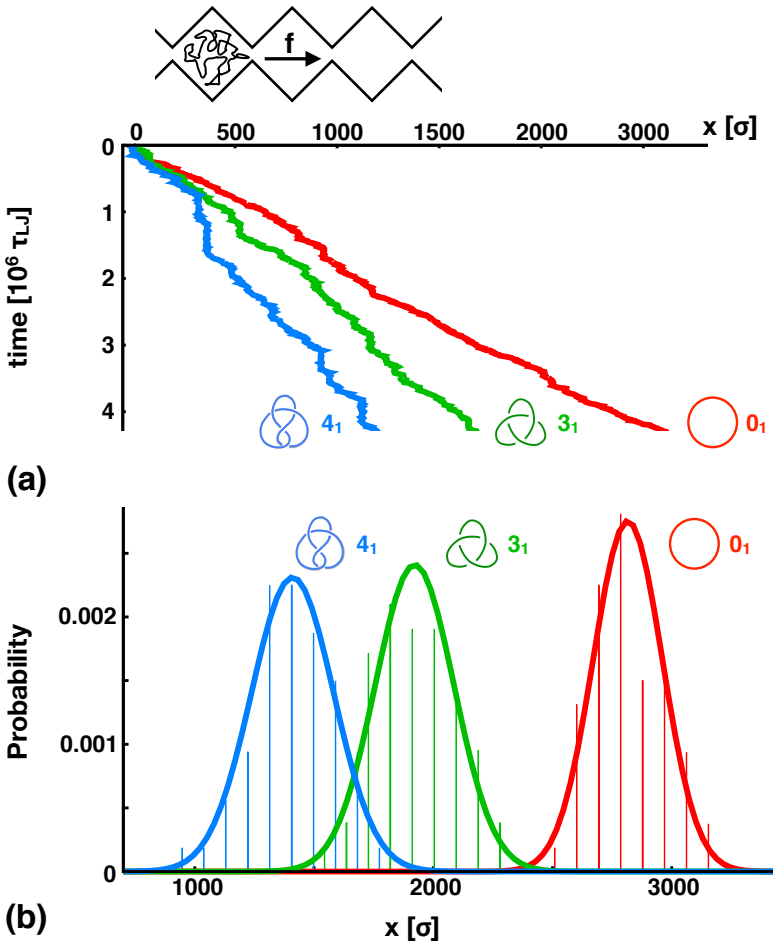


Fig. 3.18: **(a)** Typical trajectories when a driving force, $f = 0.0007\ \epsilon/\sigma$, is applied to rings with 680 beads and different topologies inside the optimally-modulated nanochannel ($L_{\parallel} = 2R_{max} = 40\sigma$, $R_{min} = 8\sigma$). **(b)** The vertical bars are the normalised histograms of the driven displacements at time lag $\tau = 4.8 \cdot 10^6 \tau_{LJ}$ collected over at least 50 independent trajectories per topology. The solid curves are Gaussians with the same mean and variance of the displacement data. Adapted from [21].

The above results refer to a passive sorting mode based on free diffusion. However, one might expect them to hold for active sorting modes, too, i.e. when translocation is driven by an external field.
Based on the simple analogy of treating the rings as single diffusive particles, one expects that, for a given driving force, f, the mobility μ is proportional to

the diffusion coefficient:

$$\langle x(t)\rangle_{t\to\infty} \simeq \mu f t = \frac{f}{\gamma}t = \frac{Df}{K_B T}t \tag{3.25}$$

The force f should not be too large, otherwise the linear-response relationship between chain mobility and diffusivity is obliterated. Yet it should also be not too low as otherwise the inherent stochasticity of the process may cause strong overlaps of the displacement distributions for different knot types.
An estimate for the crossover force between these two regimes can be obtained by neglecting the impact of pore size and considering the single remaining characteristic force of the system, $\bar{f}$, which is obtained by dividing the thermal energy $K_B T$ by the average longitudinal span of the chain. The latter is about equal to 45σ for unknotted chains of $N = 680$ beads inside the larger channel, and therefore $\bar{f} \sim 0.02\ \epsilon/\sigma$.
We verified that, when the force applied externally to each bead is much smaller than $\bar{f}$, then the displacement distributions at a given elapsed time can, indeed, be separated for the different knot types. This is shown in Fig. 3.18 where the driving force is $f = 0.0007\ \epsilon/\sigma$. In fact, the distance of any two peaks is at least 1.5 times larger than the sum of the widths of the associated distributions. The separation between unknotted rings and the collective set of knotted ones is sharp. This is a desired feature because it gives robustness to the primary distinction into trivial and non-trivial topologies. The differences are smaller, though still well detectable between non-trivial topologies. This is consistent with the progressive loss of resolution with knot complexity that is observed in gel electrophoretic experiments, which still allows to retain a good control on the abundance of knots within a certain complexity range [20]. In the current setup, the expected degradation of resolution can be counteracted, for instance, by tuning the pore size so that its competition with the size of the knot can maximally enhance the dynamic range of mobilities versus knot type.
A further interesting related point regards the variations in position and length that the knotted region experiences during the driven translocation. We found that when the ring is mostly inside one chamber, the knot is typically large and, for the considered range of N, spans a good fraction of the chain. The translocation of the ring across neighbouring chambers is, instead, accompanied by a shrinking of the knotted portion which, in the process, can also slide along the contour length. This process is shown in Fig. 3.19 for a 4_1-knotted chain that performs three consecutive jumps.

3.10 Summary and perspectives

Our theoretical and computational results illustrate the general viability of employing spatially modulated nanochannels to sort polymer rings by knot type. This is made possible by the unexpectedly rich dependence of the longitudinal mobility on the characteristic sizes of the chain, the knot and the modulated channel. Furthermore, the interplay of these length scales is aptly interpreted with the Fick-Jacobs theoretical framework. This could therefore be used as

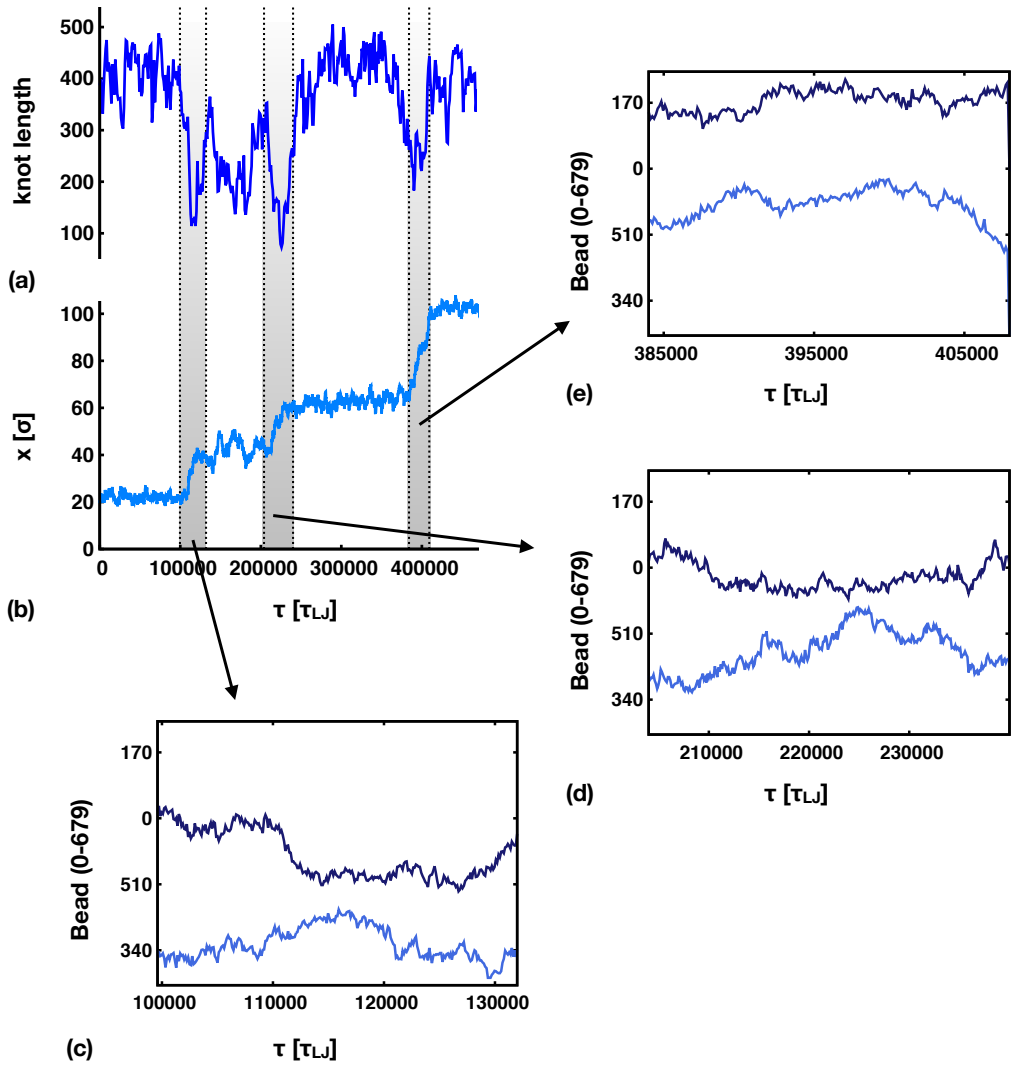


Fig. 3.19: Time evolution of the knot length **(a)** and trajectory of the polymer's centre of mass **(b)** for a 4_1 knot of $N = 680$ beads driven across the modulated channel by a force $f = 0.0007\ \epsilon/\sigma$. The parameters of the modulated channel are: $R_{min} = 8\sigma$, $L_{\parallel} = 2R_{max} = 40\sigma$. Shaded regions are the rime intervals during which the ring translocates across adjacent chambers. For these intervals, the position of the knot ends along the ring contour is shown in panels **(c)-(d)-(e)**. The position of the knot ends wraps periodically in the $0 - 679$ range. Taken from [21].

a predictive scheme to design channels optimally suited for discriminating the topologies of interest by either passive (free diffusion) or active (driving fluid or field) means. Here, we focused on the mobility of the simplest prime knots (3_1 and 4_1) within a specific modulated geometry. A possible natural development of this study would be the analysis of more complex knots, including composite ones which overwhelm prime knots with increasing chain length, and a more systematic exploration of the modulation and geometry of the confining chambers.

Note that, for simplicity and computational convenience, hydrodynamic effects in bulk and inside the channel were neglected.

A collaborative project to introduce precisely hydrodynamics effects is currently

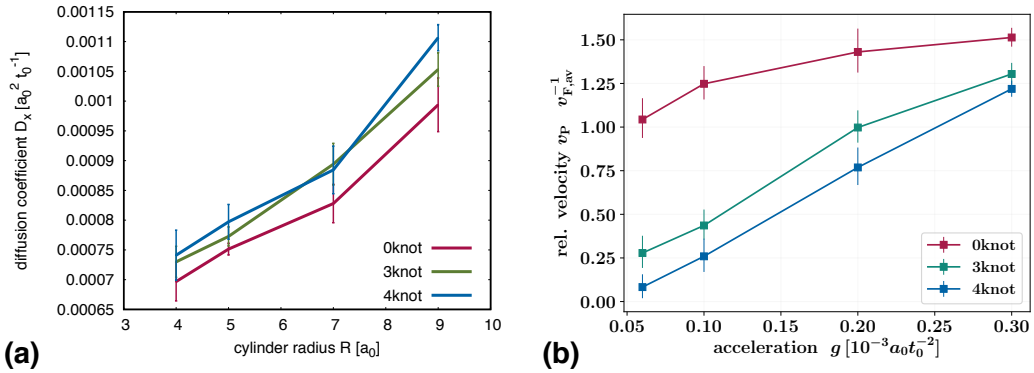
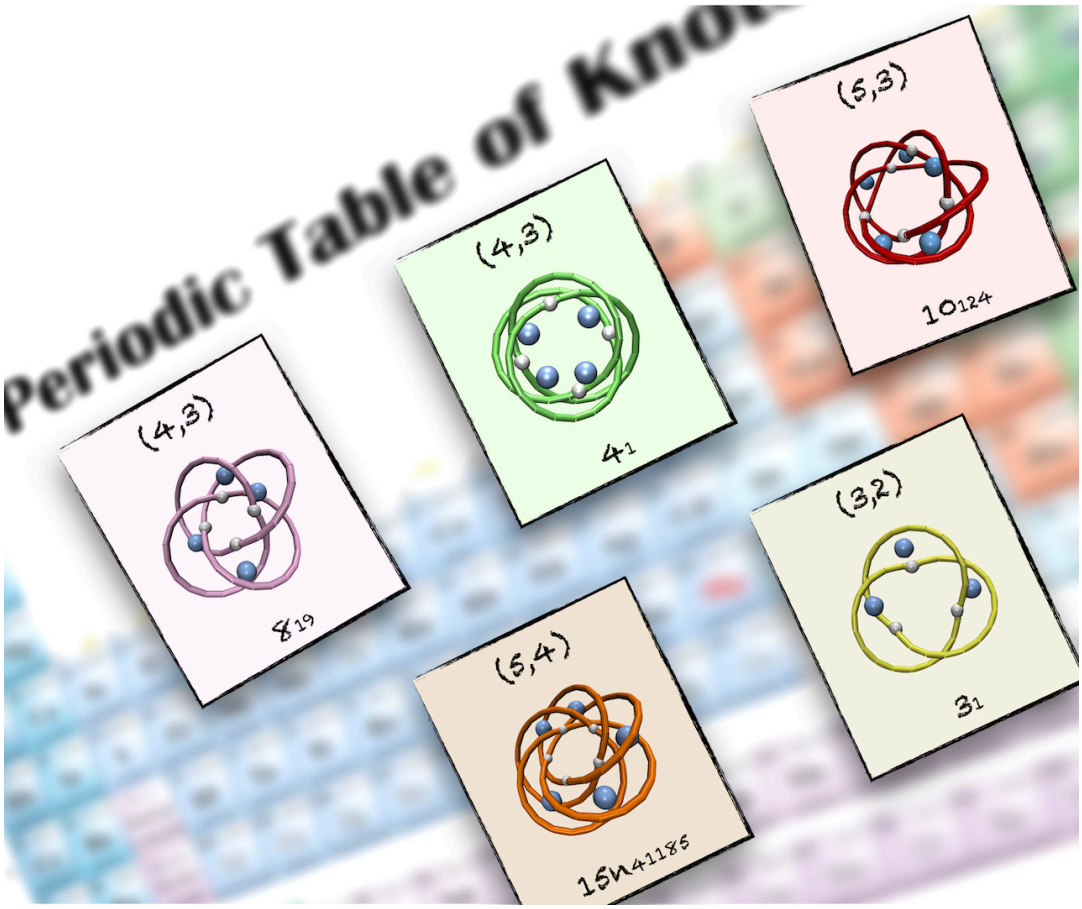


Fig. 3.20: **(a)** Diffusion coefficient in uniform channel as a function of the channel radius R for rings with $N = 150$ beads and three different topologies: 0_1 in red, 3_1 in green and 4_1 in blue. **(b)** Polymer transport velocity v_P relative to the fluid transport velocity v_F of the same chains inside a spatially modulated nanochannel as a function of the acceleration g used to create the flow. The geometry of the channel is given by $R_{min} = 4\ a_0$, $R_{max} = L_{\parallel}/2 = 9\ a_0$. Quantities are expressed in MPC units.

underway in collaboration with Ms. Lisa B. Weiss and Prof. Christos N. Likos of the University of Wien [97].
In their scheme, hydrodynamics effects have been introduced using an MPC algorithm [98] and preliminary results have shown that an intrinsic difference in diffusivity between topologies emerges even in the uniform channel case, see Fig. 3.20a. Results have shown that knotted chains diffuse faster than their unknotted counterparts for different values of the channel radius.
Further investigations were performed in case of motion under flow in optimised modulated channels, where the flow was created with a constant acceleration g on the fluid particles. Results, shown in Fig. 3.20b, illustrate that the polymer transport velocity v_P is now ranked inversely with respect to the topological complexity and that a suitable interval of g values allows the topological sorting (for the simplest knots).
We thus envisage that the proposed set up could be useful in practical applications, particularly for viral DNA molecules extracted from viral capsids [19, 69, 99] that, when they are sufficiently long (such as the λ-phage genome) are beyond reach of current electrophoretic knot sorting techniques. A further possible extension might involve the use of modulated nanoslits in place of modulated nanochannels for topological sorting [100].

Chapter 4

Discovering privileged topologies of molecular knots with self-assembly models



4.1 Background and motivation

The past two decades have seen major efforts in understanding the principles of molecular self-assembly [101–107] and in particular of the assembly of molecules with complex topologies [26, 27, 108, 109]. It has been shown that molecular knots and links have practical applications. They could be used both for anion

sensing and assisted anion catalysis [110–112], or as scaffolds to be functionalised [113–116]. Furthermore, their entangled structures could be useful as seeds for the assembly of weaving structures [117–119] and molecular nanocages [120, 121].

Molecular knots caught researchers' interest because they are the results of a set of revolutionary techniques of chemical synthesis [122], whose developers were also awarded the Chemistry Nobel Prize in 2016 [28]. In these approaches, chemical reactions between small building blocks are indirectly controlled to obtain a molecule with *a priori* target topology. However, several challenges are encountered in designing and then obtaining knotted molecular architectures and these challenges lie in the numerous concurrent physico-chemical mechanisms that need to be balanced and steered. These include, for instance, the length, thickness, curvature and relative spatial orientation of the building blocks. They are all crucial for the generation of the correct crossing points, see the molecular knots in Fig. 4.1. Furthermore, the choice of the target topologies is probably one of the major challenges in the knot synthesis [123–128], in fact not all geometrical representations are expected to be equally designable in practice [29].

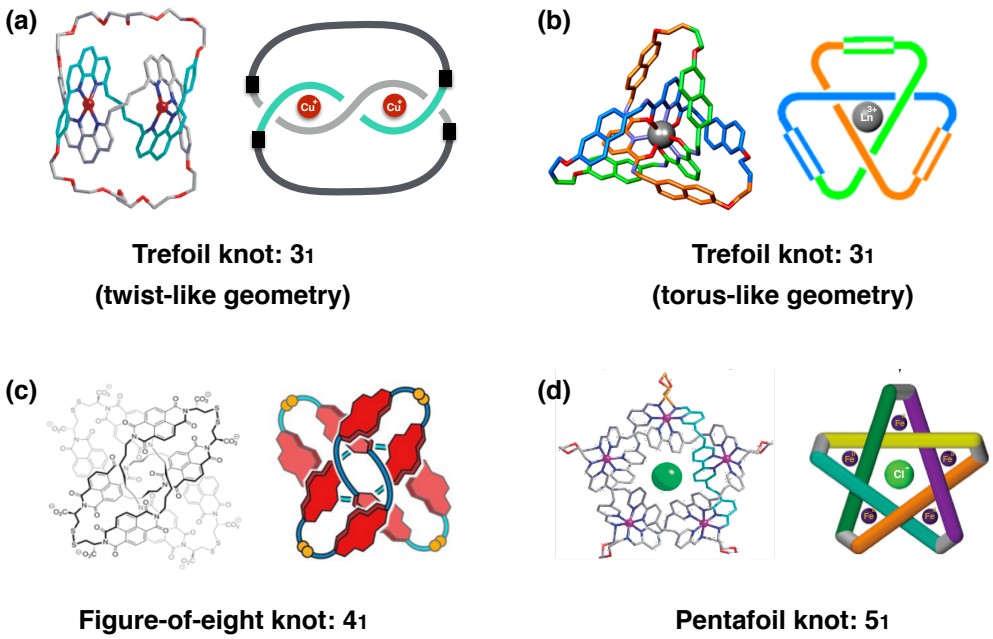


Fig. 4.1: Examples of molecular knots: the trefoil knot in **(a)** its twisted-like [26] and **(b)** torus-like geometry [129], **(c)** the figure-of-eight knot [27] and **(d)** the pentafoil knot [130].

Because of these challenges, only a handful of different topologies have been synthesised so far. Until recently [131, 132], among prime knots, only the trefoil [133], the pentafoil [130] and the figure-of-eight knots [27] had been synthesised, as well as the simplest link types [126, 134–137].

The 3_1 (trefoil) knot was the first one to be assembled. It was primary made in its twisted representation by Sauvage and co-workers in 1989 by using the metal ion templating technique, see Fig. 4.1a. In this approach, helical building

blocks were coordinated by $Cu(I)$ ions to form a double-strand linear helicate. The linear construct was then ligated into a closed topology by using suitable short ligands to obtain the desired knot [133].
Several years later, the 5_1 (pentafoil) knot was assembled with a similar technique by Leigh and co-workers. The major breakthrough was given by the assembly of a cyclic double-stranded metal helicate instead of a linear one, as shown in Fig. 4.1d. Suitable building blocks and $Fe(II)$ metal ions were crucial for the knot formation, as well as the introduction of a central anion with the role of selecting the formation of the cyclic pentafoil construct [130]. The same technique allowed also the assembly of the 3_1 knot in its torus-like geometry [129], as shown in Fig. 4.1b.
Differently from the above mentioned knots, the 4_1 (figure-of-eight) knot was created with a rather different technique by Sanders and co-workers in 2014. Instead of a metal helicate synthesis, an assembly in water of flexible building blocks was performed. Here, the hydrophobic effects on the building blocks provided an alternative driving force to promote the formation of knots [27]. This topology is shown in Fig. 4.1c.
From this brief overview, it is clear that different design strategies have been used for building knotted molecular constructs. Nonetheless, independently of the specific techniques that have been used, all the prime knots that have been produced so far, see Fig. 4.1, have two common features: they are quasi-planar and symmetric for discrete rotations around an axis perpendicular to the molecular "plane".
Complementary investigations to identify novel addressable topologies in the the synthesis of molecular knots, comes from theoretical and computation models.

Fig. 4.2: Computational prediction [138] (left) and synthetic realisation [131] (right) of the 8_{19} knot.

In this respect, a previous study of our group used simplified models to investigate the spontaneous self-assembly of knotted and linked constructs [138, 139]. This approach revealed the spontaneous assembly of a complex but designable eight-crossing knot, the 8_{19} torus knot, see Fig. 4.2. This topology was therefore indicated as a promising target for molecular synthesis. Two years after this prediction, Leigh and co-workers have precisely succeeded in creating the 8_{19} knot with the same geometry [131].

This result underscores the predictive power of general models of self-assembly and stimulated us to explore which further designable knots could be proposed as possible targets for supramolecular assembly.
In this Chapter, whose content is mainly based on the publication of ref. [30], we will use self-assembly simulations of rigid helical templates for a systematic survey of accessible knot types that show the aforementioned distinctive geometric features found in synthetic molecular knots, namely cyclic-symmetry and quasi-planarity.

4.2 The helical building blocks

In this work, we use rigid helical fragments as building blocks to investigate the repertoire of designable knotted structures.
The helical shape is actually modelled after helicates, that are commonly used in the synthesis of molecular knots [131, 133, 136, 140]. Furthermore, helical fragments can be easily described in terms of only two geometrical parameters: the projected angle α and the vertical span h, as shown in Fig. 4.3. The helical trace of fragments of unit radius is represented by the parametric equations:

$$\begin{cases} x(t) = \cos(\alpha t) \\ y(t) = \sin(\alpha t) \\ z(t) = \chi h t \end{cases} \tag{4.1}$$

where χ represents templates chirality: *right* for $\chi = +1$ or *left* for $\chi = -1$. The parameter t takes values in the range $[0, 1]$.
In our study, we discretised helical fragments of unit radius as rigid chains of beads [138] of nominal diameter $\sigma = 1/3$ (the unit of length being the helical radius), as shown in the second panel of Fig. 4.3. The beads centres are equispaced at a distance in the $[\sigma, 2^{1/6}\sigma]$ range on the helix centreline; the precise value depends on the integer number of beads necessary to cover the helical contour.

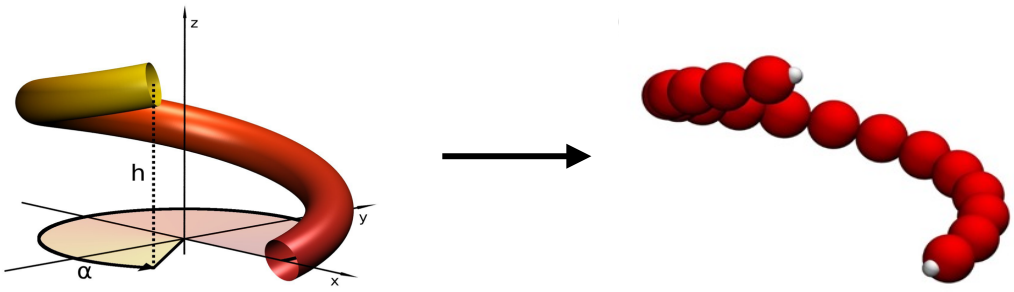


Fig. 4.3: Representation of the centreline of the helical fragment and its practical implementation as a rigid chain of beads (in red) with patchy ends (in white). This image is from ref. [138].

Helical templates are then functionalised with two small patches lying on the

surface of terminal beads at the intersection with the helical centreline (white spheres in Fig. 4.3). The attractive patches serve to join different helical fragments and therefore lead to the formation of closed complex structures.
In the next sections, we will present details and results of our computational survey. We will first use a Monte Carlo sampling technique and then molecular dynamics simulations.

4.3 Monte Carlo exploration of addressable topologies

For the Monte Carlo survey, we proceeded in two steps. First, we explored the repertoire of all topologies that are accessible to closed constructs made of a fixed number of helical templates. Then, to identify the candidate shapes of feasible realisation, we selected *a posteriori* those having the signature features of existing synthetic molecular knots: cyclic symmetry and quasi-planarity.

4.3.1 Monte Carlo method

The Monte Carlo procedure was used to explore the conformational space of closed constructs made of n_T templates. These constructs are formed by identical rigid templates, i.e., same (h, α) parameters, and are joined by exactly superposing the patches at their termini in a circular fashion, as shown Fig. 4.4.
The ensemble of possible conformations is sampled by applying unrestricted crankshaft moves that are hinged at randomly chosen pairs of the superposed patches. These moves permit to maintain the connectivity of the templates and the spatial superposition of the patches at all times (as a consequence, no patch-patch attractive potential was introduced).
All generated configurations are accepted, except for those with steric clashes (overlapping beads of different templates) or with large ($> \pi/4$) contacting angles between consecutive templates, which are rejected.
We explored ensembles of closed constructs made of $n_T = 3$, 4, or 5 helical templates, whose angular span and pitch vary respectively in the ranges $1.4\ \pi \leq \alpha \leq 1.9\ \pi$ and $0.1 \leq h \leq 2.0$. All possible combinations of right and left-handed chiralities were also considered.
For each combination of number of templates, (h, α) geometry, and chirality, we sampled 1000 configurations spaced by $5 \cdot 10^4$ Monte Carlo (crankshaft) moves, a timespan larger than the typical autocorrelation time for the considered systems, as shown in Fig. 4.5.
From the set of sampled conformations we then selected *a posteriori* those, if any, with approximate cyclic symmetry. To this end, a symmetry score was computed as the root mean square deviation (RMSD) of the structural alignment of a construct with its circular permutant with the best structural alignment [141]. The best alignment was searched over all cyclic permutations of the beads indices with an indexing shift at least equal to half the templates' length (number

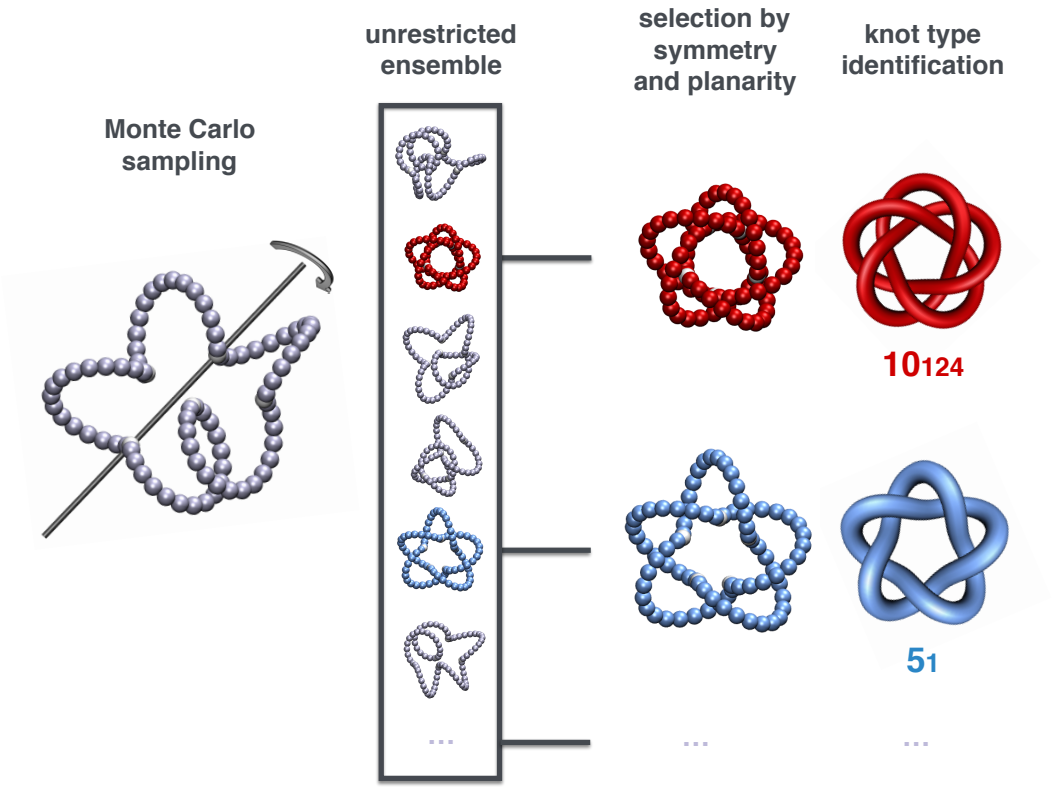


Fig. 4.4: The two-tier strategy to detect admissible topologies consists of an initial Monte Carlo sampling of closed constructs made of $n_T = 3, 4, 5$ helical templates followed by an a posteriori selection of the quasi-planar, cyclic-symmetric and knotted instances. In this $n_T = 5$ example, the ensemble generated with unrestricted crankshaft moves includes two such instances: the 10_{124} and the 5_1 knots. For visual clarity, here and in other figures the attractive patches joining the helical templates (discretised as chains of beads) are shown as white beads larger than their actual size. Taken from [30].

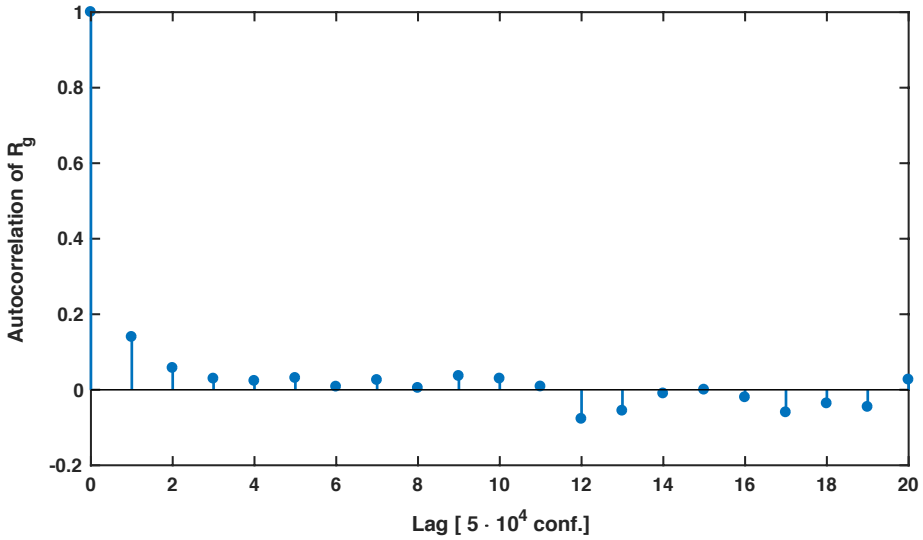


Fig. 4.5: Autocorrelation of the gyration radius of Monte Carlo sampled configurations (spaced by $5 \cdot 10^4$ moves). We considered a simulation involving closed constructs with $n_T = 5$ templates, which are all right-handed and have geometry ($h = 0.4, \alpha = 1.8\pi$).

of beads).
Symmetric conformations were found for equal chirality templates and, in case of $n_T = 4$, for a mixture of two left- and two right-handed templates. For $n_T = 3$ and 4 templates, the presence of cyclic-symmetric constructs is signalled by a peak or shoulder at low values of the symmetry score distribution. For five templates, no peak is discernible and therefore we took the RMSD cutoff value as the largest RMSD below which all constructs are cyclic-symmetric. In Fig. 4.6, we show distributions and respective cutoffs of the symmetry score.

4.3.2 Knot analysis

As introduced in Chapter 1, here and in the next sections we combined two methods to identify knot types. For most constructs, the projected patterns of crossings was simple enough that the Alexander polynomial sufficied to identify the knot type [33]. For more complicated patterns, the Dowker code was compared against a look-up table for prime knots of up to 16 crossings using the Knotscape software package [42]. For knots that could not be simplified below 16 crossings and that were not factorisable into simpler components, we used the Dowker Code as the topological identifier.

4.3.3 Monte Carlo results

The results of the Monte Carlo discovery procedure are summarised in Fig. 4.7. The figure presents the repertoire of the different types of closed and cyclic-symmetric knotted constructs found across the entire explored range of shapes and chiralities of the building blocks.
In particular, the graphical table displays the structural representatives of each knot type. Strikingly these are all inherently oblate, meaning that quasi-planarity was co-opted by the cyclic-symmetry constraint and did not need to be enforced additionally.
Each representative in Fig. 4.7 is accompanied by a number of key properties, namely the braid representation of the knot (in linear, circular and symbolic form) and the topological designability score. We define the latter as the number of distinct template shapes, i.e. distinct points in the discretised (h, α) space, that can assemble in that particular cyclic knot. This measure is reflective of the designability of the knot, that is the robustness to variations in template shapes. This ought to be a valuable quantitative element for guiding the design of novel molecular topologies. The regions in the (h, α) parameter space where the various knot types occur, are highlighted in the topological state diagram of Figs. 4.8 and 4.9 .
The main result of Fig. 4.7 is that, across the wide range of template shapes and combinatorics of sampled constructs, the phenomenological selection is survived by only few privileged topologies.
These cover all knot types that have been synthesised so far and, above all, include novel ones too. Entries in Fig. 4.7 that correspond to known molecular knots are: (i) the 3_1 knot in the two classic geometrical variants: the ideal-like and the twisted one respectively obtained with 3 and 4 templates, as in the

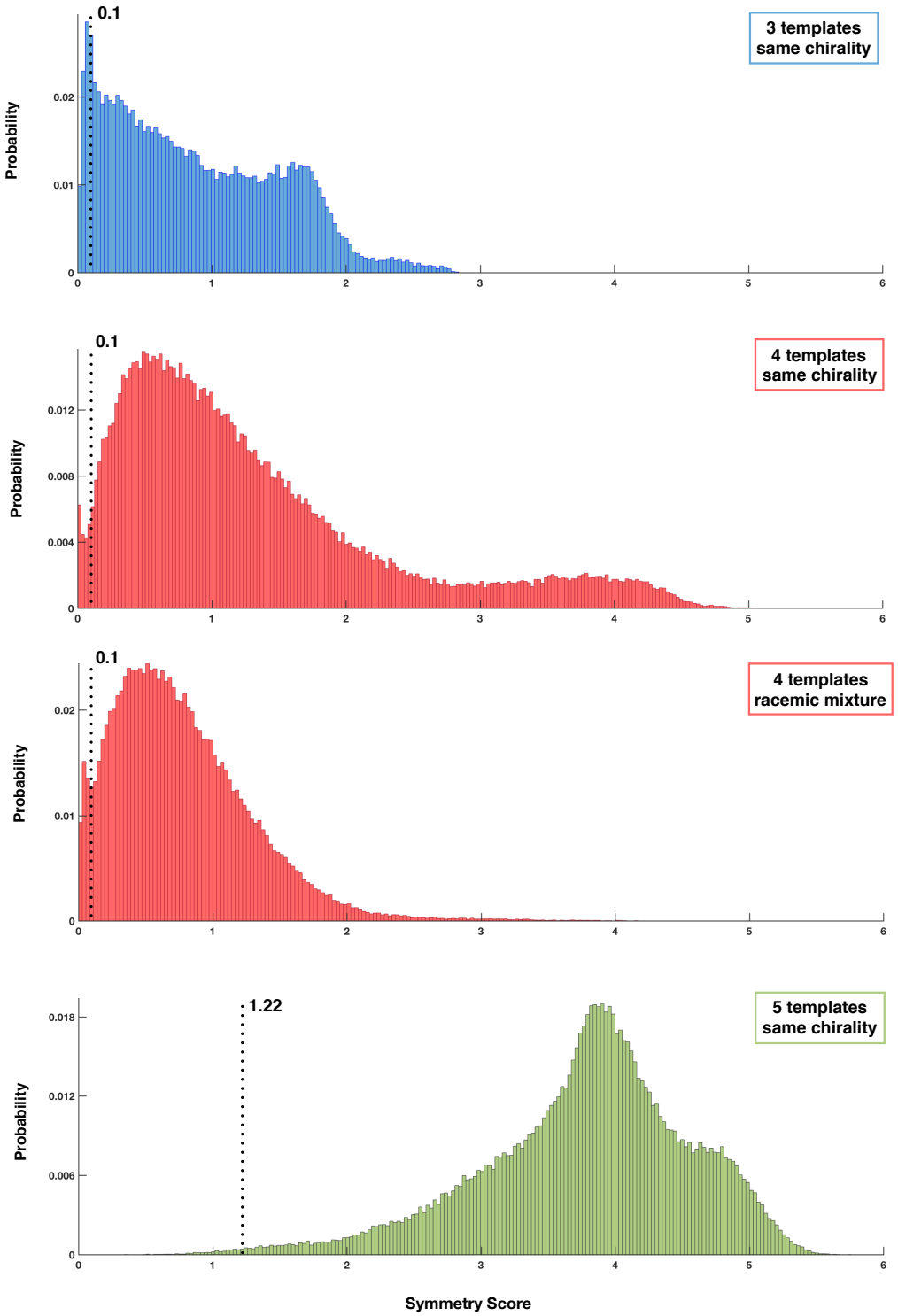


Fig. 4.6: Probability distribution of the cyclic-symmetry score for Monte-Carlo sampled constructs made of $n_T = 3$, 4, and 5 templates of same or different chiralities, as indicated. Each distribution is cumulated over all considered templates' shapes. The cutoff value for the score used to select such instances is marked with a dashed line. Taken from [30].

experimental synthetic realisations (Fig. 4.1a and b) [129, 133]; (ii) the 5_1 knot made of 5 templates, as in the molecular realisation of Fig. 4.1d [130] and (iv)

the aforementioned 8_{19} knot made of 4 templates, again as in the recent study of Fig. 4.2 [131]. The instances are all torus knots [31].
The only non-torus knot in Fig. 4.7 is the 4_1, or figure-of-eight knot. Interestingly, one notes that this amphichiral twist knot can be established with four templates, either with the same or opposite chirality. The latter, racemic combination is much favoured because it covers a significantly wider region of parameter space, see designability score and Figs. 4.8 and 4.9 , and yields more planar constructs. The 4_1 topology has been experimentally obtained before, too, and with the same number of templates, though the building blocks were not helicates but flexible diblock modules (Fig. 4.1c) [27]. The 4_1 instance in Fig. 4.7 therefore makes the important point that this topology would be realisable with non negligible probability also by using metal templating technique

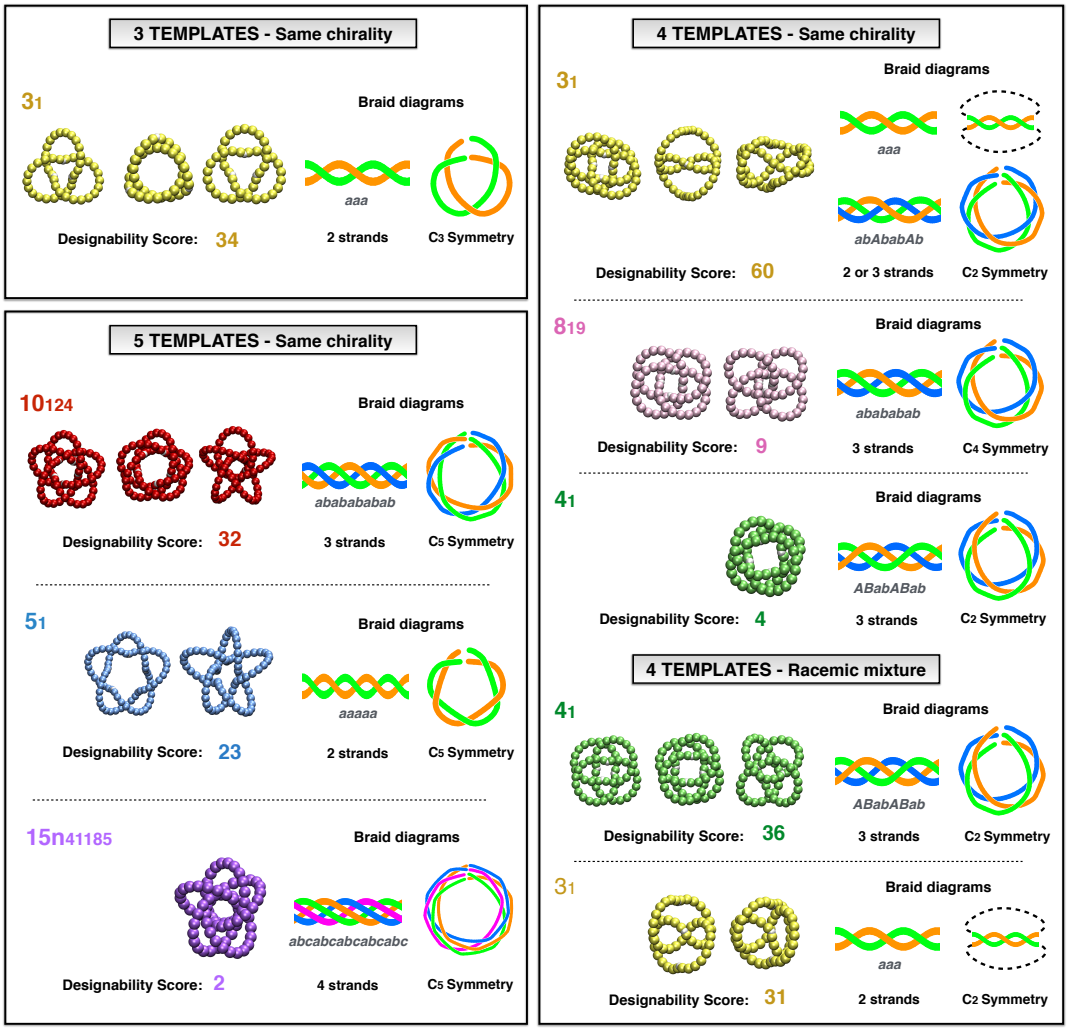


Fig. 4.7: The graphical table provides the complete repertoire of non-trivial topologies that can be realised with up to 5 identical helical templates (of same or opposite chirality) in the form of cyclic-symmetric and quasi-planar constructs. For each knot type we show one or more representative conformers (grouped by the number of templates) and the corresponding linear, circular and symbolic braid representations. The designability score is the number of points in the discretised (h, α) parameter space of helical template shapes for which that knot type is observed. Adapted from [30].

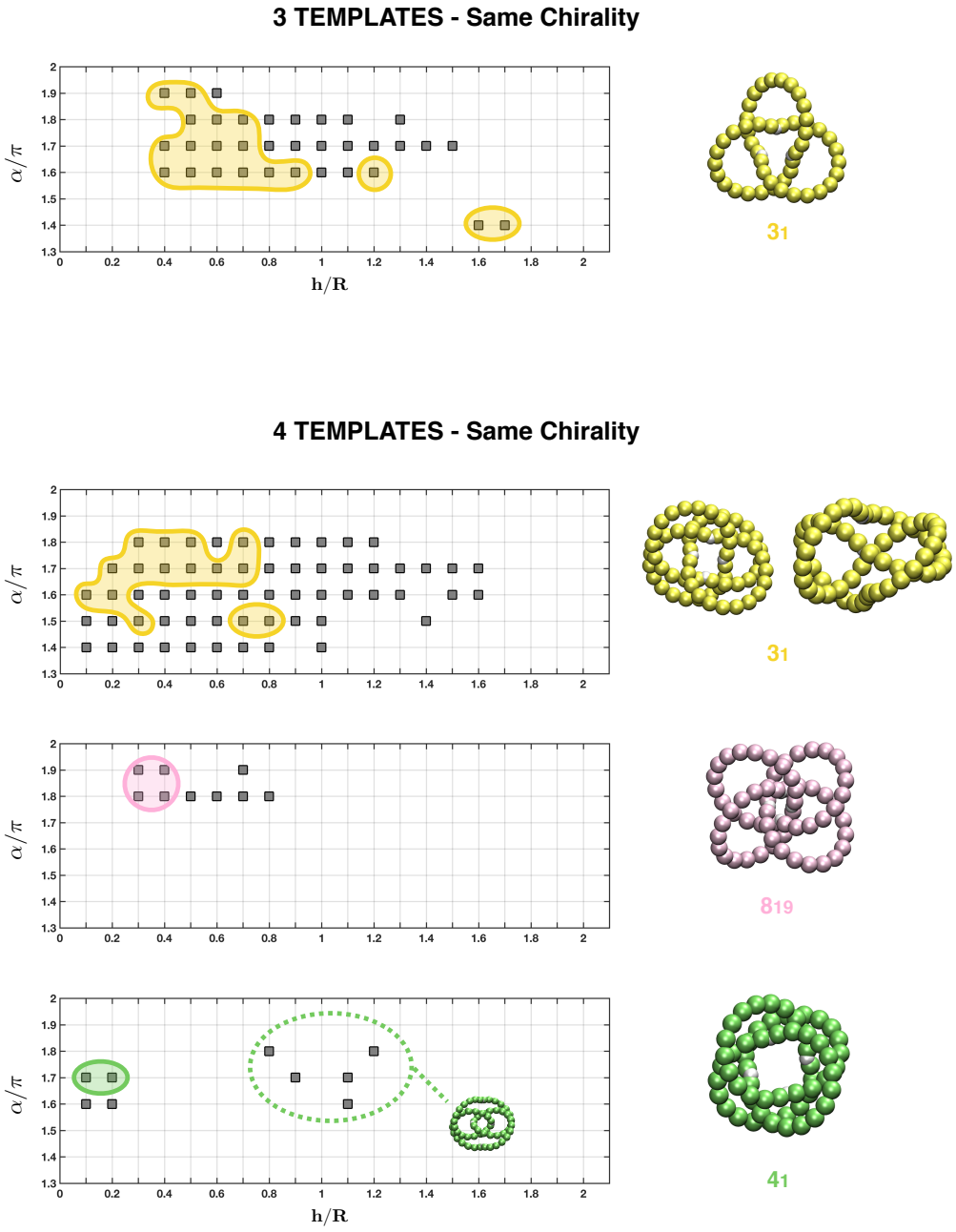


Fig. 4.8: Topological state diagram. The highlighted points in the discretised (h, α) parameter space mark the template shapes where one observes constructs $n_T = 3$ (top) and $n_T = 4$ (bottom) templates with the topologies sketched on the right. The highlighted regions include points where the indicated knots account for more than 1% of the MC-sampled constructs. The cases shown here are made of templates with the same chirality. The 4_1-knotted instances in the bottom panel mostly correspond to the geometry shown on the right, but there are also instances of an alternative geometry, shown inside the topological state diagram. Taken from [30].

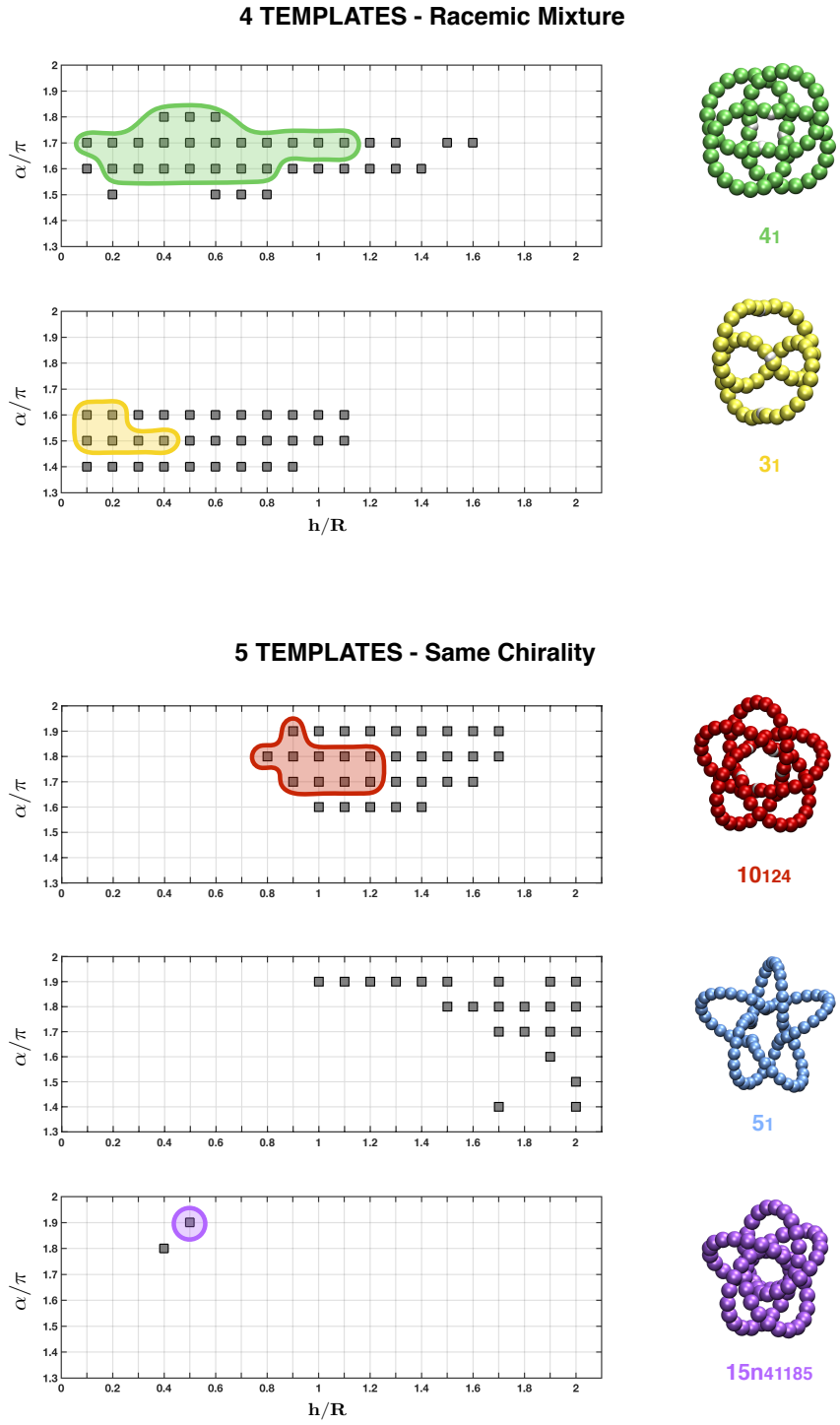


Fig. 4.9: Topological state diagram. The highlighted points in the discretised (h, α) parameter space mark the template shapes where one observes constructs $n_T = 4$ (top, racemic template mixture) and $n_T = 5$ (bottom, templates of same chirality) templates with the topologies sketched on the right. The highlighted regions includes points where the indicated knots account for more than 1% of the MC-sampled constructs. Taken from [30].

and helicates, but the latter should ideally be of opposite chirality. This is a features that, to our knowledge, has not yet been systematically explored. Note that a further symmetric geometry for the 4_1 knot is possible, as shown in Fig. 4.8.
The fact that none of the molecular prime knots realised before the completion of this study, are missing from the graphical table, gives confidence in the feasibility to synthesise in the future the two novel topologies appearing in Fig. 4.7. These are the 10_{124} and $15n_{41185}$ knots. Though being made by only $n_T = 5$ templates (same chirality), like the 5_1 synthetic knot, these topologies are more complex than the previous mentioned synthetic knots obtained so far. In fact, it is arresting to observe that, out of the millions of prime knots with up to 15 crossings, those feasible with a handful of cyclically-arranged templates are only 6.
The 10_{124} topology emerges as a particularly designable cyclic knot, because it can be obtained with tens of different template shapes. It is intriguing that this topology is among those that are only rarely populated in aspecific self-assemblies [138], i.e. without cyclic symmetry or quasi-planarity, and appears, with other instances, in energy-minimising arrangements of dipolar particles [142].
The 15-crossing knot is an even more remarkable example of sophisticated interplay of geometry and topology, with as many as 15 crossings established by 5 templates only. It forms abundantly for a specific template shape, the one involved in the construct shown in Fig. 4.7, and more limitedly for a second template geometry, see Fig. 4.9 .
Of these two novel complex topologies, we therefore single out the 10_{124} knot as the most promising one to be realised with current self-assembly strategies.

4.4 Kinetic accessibility of self-assembling knotted structures

To test the kinetic accessibility of the topologies discovered by Monte Carlo sampling, we performed hundreds of self-assembly simulations of patchy helical templates, varying both their number and geometry.

4.4.1 Molecular dynamics method

We used Langevin molecular dynamics simulations to evolve the system of initially detached patchy helical templates. The attractive interactions between the patches on the helical fragments lead to the attachment of different templates and to the formation of closed complex constructs. Furthermore, we have also enhanced the formation of closed structures with a number of free particles in solution that do not bind to each other but can attract, and hence coordinate, two or more templates, see Fig. 4.10. The role of these coordinating particles is to mimic the effects of metal ion templating [143].

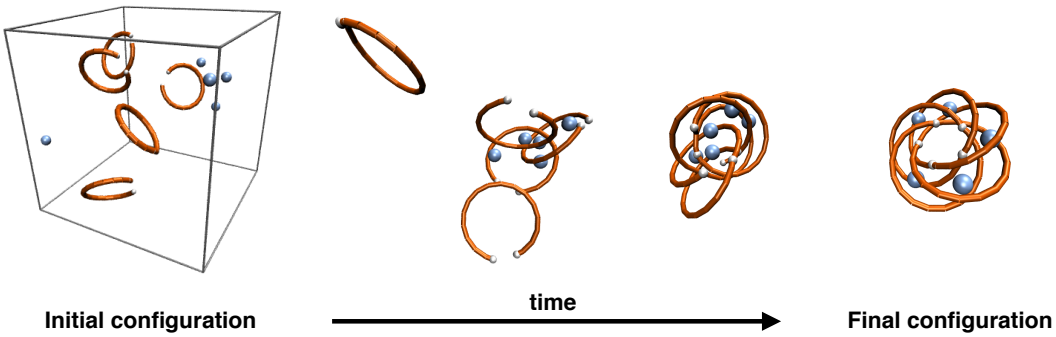


Fig. 4.10: Series of snapshots from a molecular dynamics simulation where five helical templates and an identical number of coordinating particles self-assemble spontaneously into a cyclic-symmetric $15n_{41185}$ knot. Adapted from [30].

Except for the small patches, all other particles (templates beads and the coordinating particles) interact sterically through a truncated and shifted Lennard-Jones potential:

$$U_{LJ}(d_{i,j}) = \begin{cases} 4C_{LJ}\epsilon \left[\left(\frac{\sigma}{d_{i,j}}\right)^{12} - \left(\frac{\sigma}{d_{i,j}}\right)^{6} + \frac{1}{4} \right] & \text{if } d_{i,j} < 2^{1/6}\sigma \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

where $d_{i,j}$ is the distance between the two particles. The magnitude of the potential is set to $C_{LJ} = 100$.
The patches, instead, interact via the attractive Gaussian potential:

$$U_{patchy}\left(d_{i,j}\right) = -C_p\epsilon \exp\left(-\frac{d_{i,j}^2}{2\sigma_p^2}\right) \tag{4.3}$$

Its magnitude is set $C_p = 25$ and its range $\sigma_p = 0.1\sigma$. As we will see in the following section, the combination between the patchy and Lennard-Jones interactions results in an attractive directional potential well of depth $\sim 18\ K_BT$ between helical templates.
The attraction of the coordinating particles and the templates beads is described by a Yukawa-like potential:

$$U_{Yukawa}\left(d_{i,j}\right) = -C_Y \frac{\sigma}{d_{i,j}}\epsilon \exp\left(-\frac{d_{i,j}}{l_Y}\right) \tag{4.4}$$

where $C_Y = 8$ and $l_Y = \sigma$. A repulsive version of this potential (same form but opposite sign) is introduced between the coordinating particles to keep them apart.
The interaction potentials between helical beads, coordinating particles and patches are sketched in Figure 4.11.
Langevin simulations were carried out in a periodic cubic box of size L with the LAMMPS simulation package [51]. In particular, patchy helical templates were treated as rigid bodies with LAMMPS *fix rigid* command. Default values for the beads mass and viscous friction were used, as introduced in Chapter 2, and the integration time step was set equal to $\Delta t = 0.006\ \tau_{LJ}$.

We performed simulations with a small and equal number of templates and coordinating particles, $n_T = 3, 4, 5$. Different template geometries were explored, varying the angular span and pitch respectively in the ranges $1.4\ \pi \leq \alpha \leq 1.9\ \pi$ and $0.1 \leq h \leq 2.0$, as in Monte Carlo. Combinations of templates with different chiralities were explored, too. For each value of n_T and template geometry we carried out 20 independent simulations at fixed template number density $\rho_T = n_T/L^3 = 0.0125$, where L is in units of the helical radius. The duration of each simulation was set equal to 264000 τ_{LJ}, so to be sufficiently long to observe spontaneous binding and assembling of templates, but still shorter than their typical unbinding time, which is $\sim 4000000\ \tau_{LJ}$, as shown in the next section.

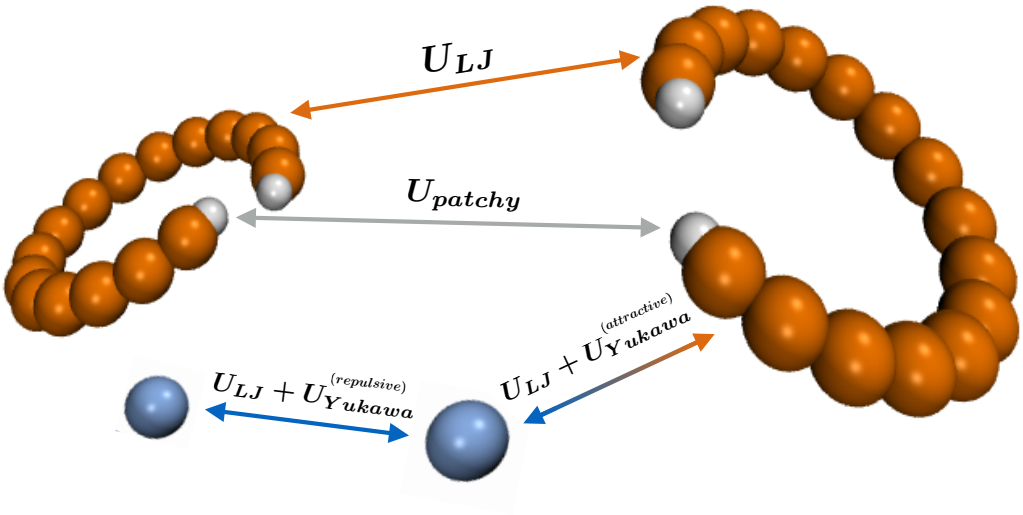


Fig. 4.11: Representation of the potentials acting between helical beads, patches and coordinating particles for molecular dynamics simulations.

4.4.2 Parametrisation of the attractive patchy interaction

The attractive patchy potential and the relative parameters were chosen as in ref. [138], where they had been tuned to obtain a pairwise directional interaction that permits the formation of stable closed constructs.
The formation of stable structures is ensured by the strength of the bonding potential that is controlled by the adimensional parameter $C_p = 25$. This choice corresponds to a depth of the Gaussian well between two patches equal to 25 K_BT. However, the effective unbonding barrier is appreciably smaller than this, as it is clarified by computing the effective free-energy profile, $F(r)$, of two patchy particles as a function of their distance r,

$$F(r) = -K_BT \log(Z(r)) \tag{4.5}$$

where $Z(r)$ is the canonical partition function integrated over the degrees of freedom, $\{\theta, \phi, \theta', \phi'\}$ defining the relative orientation of the patchy particles at

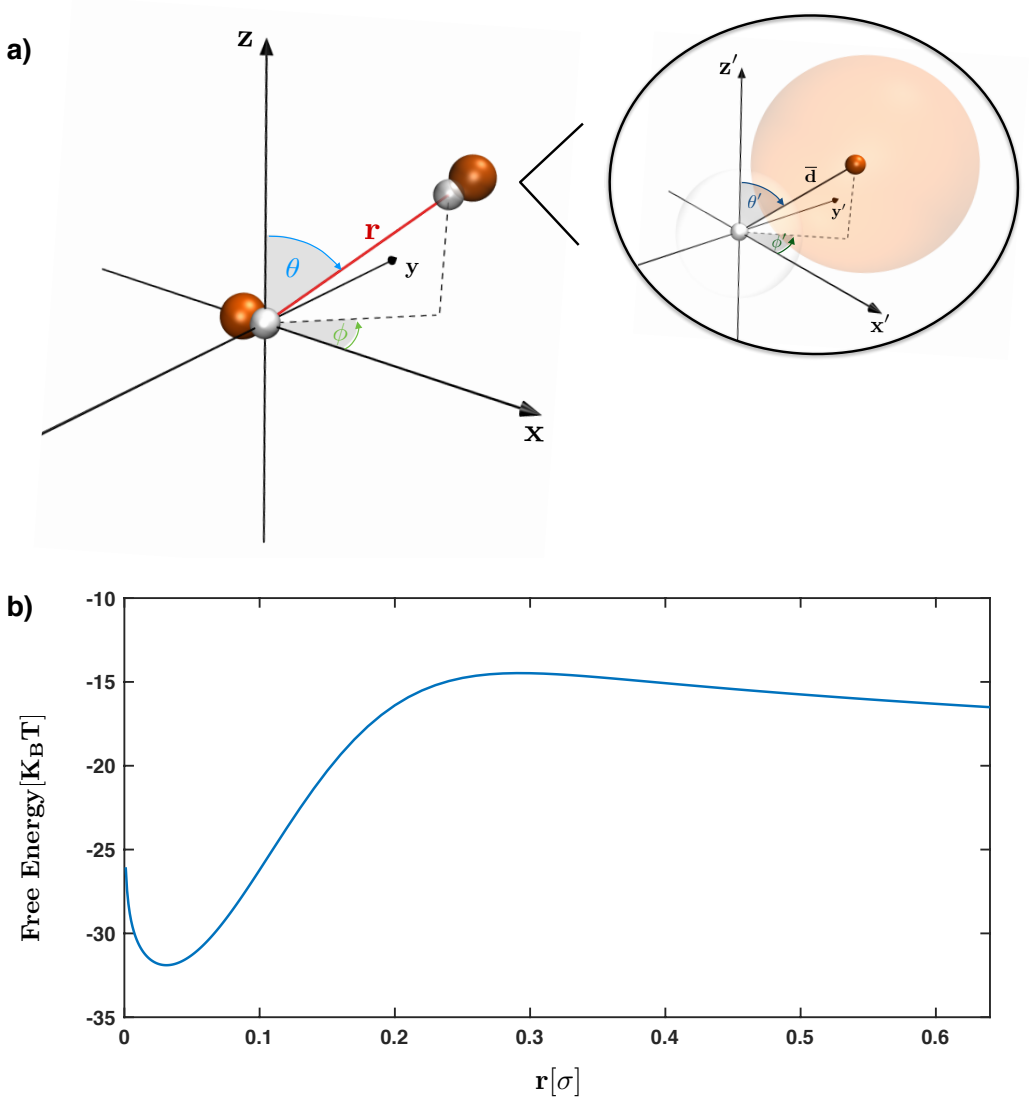


Fig. 4.12: **(a)** Representation of the degrees of freedom $(\theta, \phi, \theta', \phi')$ defining the relative orientation of two patchy particles at a distance r. **(b)** Free-energy profile, $F(r)$, obtained by numerical integration for $C_p = 25$. Taken from [30].

the given distance r. Apart from a prefactor, contributing only to an additive shift of $F(r)$, $Z(r)$ is given by:

$$Z(r) = \left(\frac{r}{\sigma}\right)^2 \int_{-1}^{+1} d\cos(\theta) \int_0^{2\pi} d\phi \int_{-1}^{+1} d\cos(\theta') \int_0^{2\pi} d\phi' e^{-[U_{patchy}(r)+U_{LJ}(r,\theta,\phi,\theta',\phi')]/K_B T} \tag{4.6}$$

where θ and ϕ are the radial and azimuthal angles of the second patch defined in the Cartesian frame centred in the first patch and with x axis oriented along the principal axis of the first patchy particle itself, see Fig. 4.12a. The other parameters, θ' and ϕ' are instead the radial and azimuthal angles that specify the orientation of the centre of the second patchy particle with respect to its patch, see inset of Fig. 4.12a.
The bead-patch distance is fixed and equal to $\overline{d} = 2^{1/6}\sigma/2$.
Numerical integration of $Z(r)$ yields (up to an additive constant) the free energy

profile shown in Fig. 4.12b.

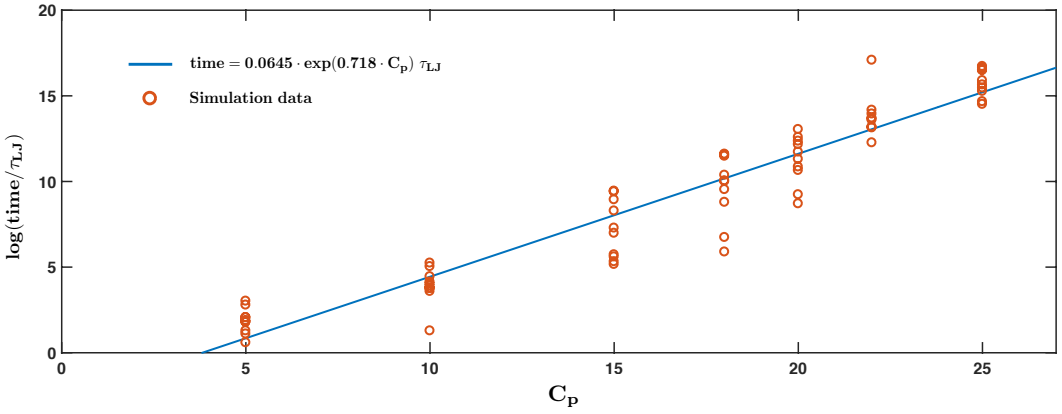


Fig. 4.13: Natural logarithm of the time required to break the bond between two patchy helical templates as a function of the strength of the patch-patch potential C_p. Results from simulations are represented with orange circles, while linear fit is represented with a continuous blue line: $time = 0.0645 \cdot \exp(0.718 \cdot C_p)\ \tau_{LJ}$. Taken from [30].

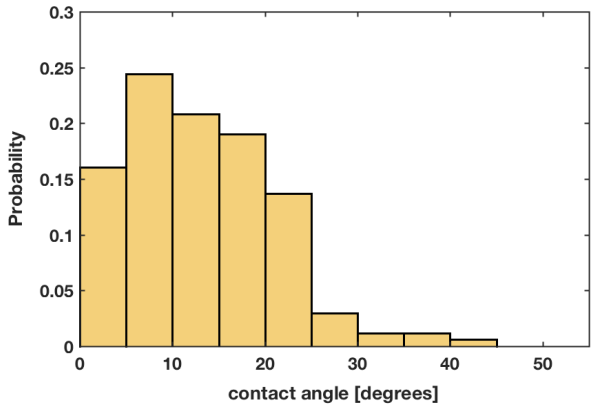


Fig. 4.14: Probability distribution of the contact angle between consecutive templates in closed constructs in molecular dynamics simulations. Adapted from [30].

One sees that due to the interplay of entropic and enthalpic effects, the effective barrier for breaking a bond is smaller than 25 K_BT, and specifically it is equal to about 18 K_BT.

The result is consistent with the actual bond-breaking kinetics for two patchy helical templates, as it is shown in Fig. 4.13. The semi-log plot shows the C_p dependence of the time required to break the bond between two initially-contacting templates during various MD simulations (ten per each C_p value). The best fit line in Fig. 4.13 is $time \propto \exp(0.718 \cdot C_p)\,\tau_{LJ}$. For $C_p = 25$, this yields the effective barrier $0.178{\cdot}25\ K_BT \sim 18 K_BT$. The associated detachment time is of the order of 4000000 τ_{LJ}, which is about 20 times larger than the typical duration of our production runs.

The directionality of the potential is essential to avoid the possible formation

of branched assembled constructs. This is ensured by a very short interaction range ($\sigma_p = 0.1$) and a very hard magnitude of helical spheres Lennard-Jones potential ($C_{LJ} = 100$). These values result in a typical contacting angle of $15° - 20°$ as shown in the probability distribution in Fig. 4.14.

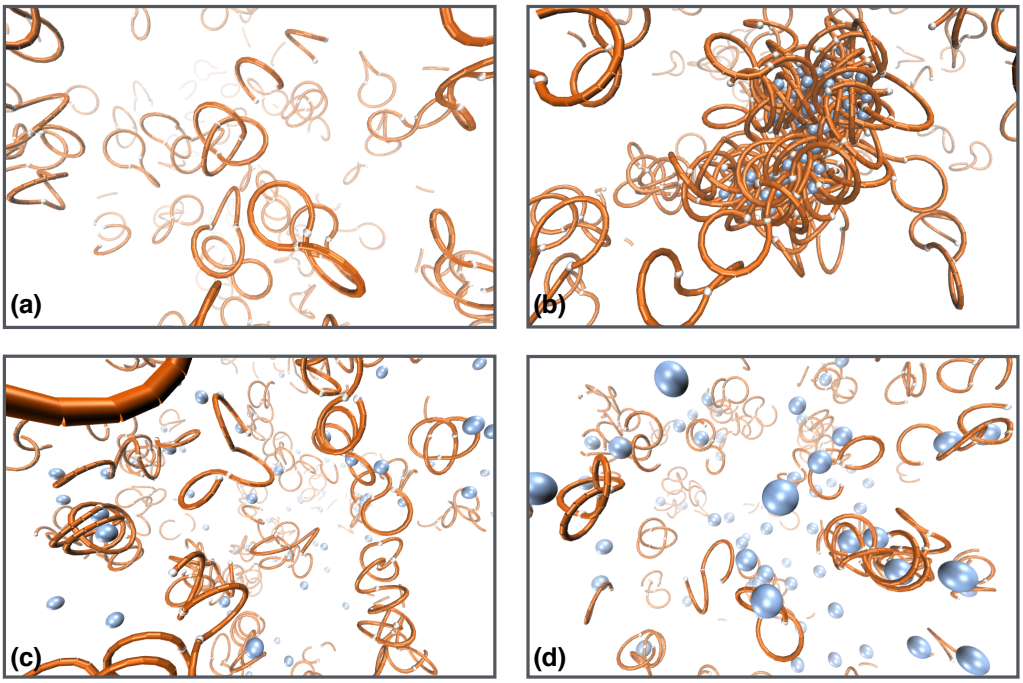


Fig. 4.15: Snapshots of molecular dynamics simulations for the self-assembly of 250 templates in various conditions. Self-assembly of templates without coordinating particles: **(a)** same chirality helical fragments with geometry $h = 1.0$ and $\alpha = 1.8\pi$ and helical fragments density 0.0125. Self-assembly of templates with coordinating particles: **(b)** racemic mixture of helical fragments (ratio 50%) with geometry $h = 0.8$ and $\alpha = 1.7\pi$, helical fragments density 0.0125, coordinating particles diameter σ and Yukawa parameters $C_Y = 8$ and $l_Y = \sigma$. **(c)** racemic mixture of helical fragments (ratio 50%) with geometry $h = 0.8$ and $\alpha = 1.7\pi$, helical fragments density 0.0125 and Yukawa parameters $C_Y = 10$ and $l_Y = 0.91\sigma$. The diameter of the coordinating particles in this example is 2σ, that is twice the size of the templates' beads. **(d)** Racemic mixture of helical fragments (ratio 50%) with geometry $h = 0.8$ and $\alpha = 1.7\pi$ and helical fragments density 0.0125; coordinating particles diameter 3σ, Yukawa parameters $C_Y = 5$ and $l_Y = 1.59\sigma$. For cases (c) and (d), the reference "contact distance" was suitably changed from the $2^{1/6}\sigma$ standard value, which is appropriate only for the default case of equally-sized template beads and coordinating particles. Other parameters are set to their default values. Taken from [30].

4.4.3 Parametrisation of the interactions between templates and coordinating particles

Various additional simulations were carried out with a wide range of number (and relative size too) of templates and coordinating particles, from few to several hundreds, see Fig. 4.15, as well as with different interaction parameters.

The simplest conditions that take to the formation of symmetric closed constructs were found with a small number of helical templates and coordinating particles. A dispersion of hundreds of templates could result in the formation of big clusters, as shown in Fig. 4.15b, and therefore in a small incidence of knots. For simulations with a small number of templates and coordinating particles, the Yukawa-like potential parameters $C_Y = 8$ and $l_Y = \sigma$, as well as the diameter σ of coordinating particles were found, through empirical observations, as suitable choices for the formation of complex knotted structures.

4.4.4 Molecular dynamics results

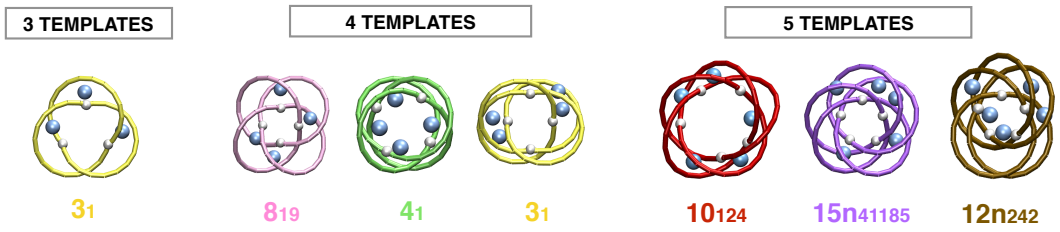


Fig. 4.16: Repertoire of cyclic-symmetric quasi-planar knotted constructs that recurrently form in self-assembly simulations. The $12n_{242}$-knotted construct is nearly cyclic-symmetric and is included for its high statistical incidence. For visual clarity the templates are represented with their centreline, omitting the constitutive beads. Adapted from [30].

As mentioned above, we performed self-assembly simulations for producing cyclic-symmetric quasi-planar constructs with systems with an equal number of templates and coordinating particles, $n_T = 3, 4, 5$ at a suitable volume density.
With these systematic self-assembly simulations, we established that all privileged topologies listed in Fig. 4.7 except for the 5_1 knot, are indeed kinetically accessible. Their representative structures formed by self-assembly are shown in Fig. 4.16, while their designability score in Tab. 4.1. These conformers include the novel complex target topologies, 10_{124} and $15n_{41185}$. This reinforces their general viability as promising targets for molecular constructs with addressable topology.

	3 templates	4 templates			5 templates		
Topology	3_1	8_{19}	4_1^*	3_1	10_{124}	$15n_{41185}$	$12n_{242}$
Designability Score	42	24	20	55	11	5	9

Tab. 4.1: Designability score for the symmetric (and quasi-symmetric) knots obtained in self-assembly of 3, 4 and 5 templates with shape parameters ($h \in [0.1, 2.0]$ and $\alpha \in [1.4, 1.9]\pi$). (*) The 4_1 is assembled from a racemic combination of 2 left- and 2 right-handed templates. Taken from [30].

The formation of symmetric 5_1 knots in MD assembly simulations is suppressed because of a large contacting angle that would be necessary to form the construct. In fact, the typical contacting angles in MD, which promote the collinearity of the contacting templates ends, are much lower than those observed in the symmetric forms of 5_1 knots of MC simulations, as shown in the probability distributions in Fig. 4.17.
Interestingly, besides the topologies of Fig. 4.7, we recurrently observed the formation of the $12n_{242}$ topology, as shown in Fig. 4.16 and Table 4.1. This knotted structure is self-assembled from five templates and differs from the $15n_{41185}$ by a localised "defect" in the otherwise regular pattern of over- and under-crossings. From its abundance, we conclude that the $12n_{242}$ topology, though not cyclic-symmetric, might be obtainable as a likely by-product of the target $15n_{41185}$ knot.

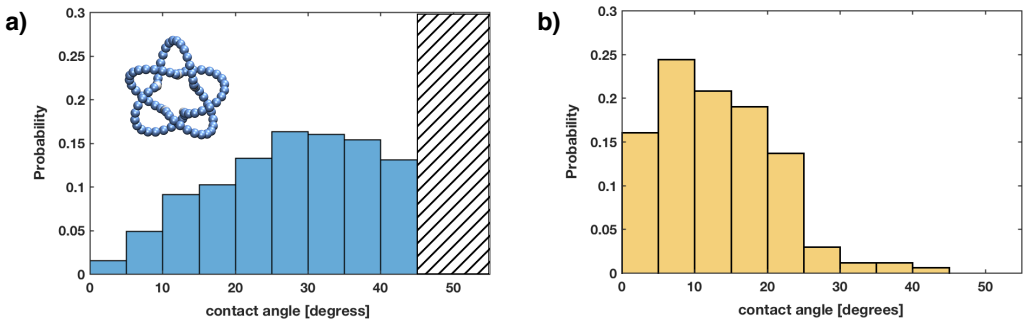


Fig. 4.17: **(a)** Probability distribution of the contact angle between consecutive templates in the Monte Carlo-generated cyclic-symmetric 5_1 knots made of 5 templates (data cumulated over all explored template shapes). Note that angles larger than $\pi/4$, or 45 degrees, are disallowed and hence are not populated. The analogous distribution, but pertaining to molecular dynamics simulations, are shown in panel **(b)**. In this case, the shape of the distribution is controlled by the steric and patchy interactions of the templates which promote the collinearity of the contacting templates' ends. Taken from [30].

4.5 Enumerative survey of cyclic-symmetric entangled structures

The repertoire of symmetric topologies of Fig. 4.7 and 4.16, shows an underlined geometrical simplicity. In fact, the braids that generate most of the knots, are characterised by the repetition of equal modular units, each of them formed by an ordered (increasing) sequence of essential braids, as shown in Fig. 4.18. The number of templates n_T matches the number of times the modular unit occurs, while the number of strands n_S is related to the number of essential braids in every unit $(n_s - 1)$. For instance, the 8_{19} knot emerges from a 4-fold repetition of the *ab* crossing pattern, while the $15n_{41185}$ knot from a 5-fold repetition of the *abc* sequence. However, if we generalise to braids that include modular units with different crossings signs, we obtain also the other possible instances of symmetric knots, like for example the 4_1 and 3_1 knots made with $n_T = 4$ and

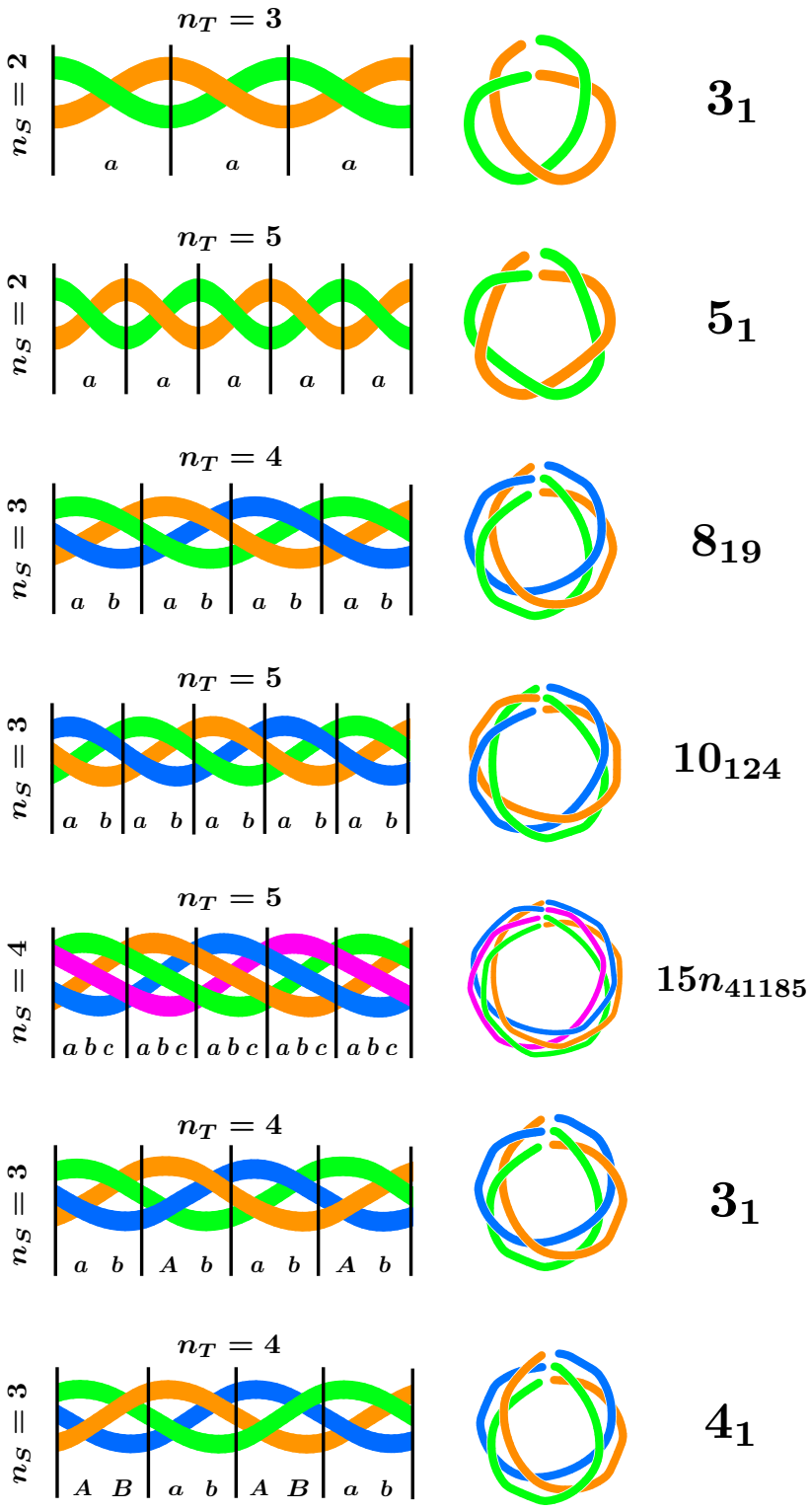


Fig. 4.18: Braid representation of the symmetric topologies resulting from our computational survey. The number of templates n_T matches the number of times the crossing pattern is repeated. The number of strands n_S is related to the number of essential braids in every repeated unit $(n_S - 1)$.

$n_S = 3$, see Fig. 4.7 and Fig. 4.18.
The geometrical and topological repertoire of Fig. 4.7 can therefore be recapitulated in terms of only two key parameters: n_T and n_S. Both parameters have, in fact, proved useful before in systematic explorations of possible synthetic topologies [131], arguably because they reflect different aspects of the practical difficulty of their realisation.
Classifying conformers in terms of n_T and n_S is an apt way to code for their structure independently of the specific geometry of their templates. For this reason, a systematic search of the (n_T, n_S) parameter space can reveal additional addressable topologies besides those in Figs. 4.7 and 4.16 made of strictly-helical building blocks or, else, rule out the existence of alternative ones.
Accordingly, we completed the survey with an exhaustive enumeration of braids, exploring all the possible crossings signs (as defined in Chapter 1), for different combinations of (n_T, n_S). For practical reasons, we limited the combinatorial search up to $n_T = 7$ and $n_S = 4$, and we assumed $n_T \geq n_S + 1$, meaning that viable templates should cover less than a full turn when projected.
The results are shown in Fig. 4.19. Note that, since the search scheme based on braiding combinatorics is more general than the Monte Carlo exploration, the topologies listed for each (n_T, n_S) pair are not limited to those in Fig. 4.7.
As mentioned above, we explored all the braids that consist of a repetition of modular units of essential crossings with all possible combinations of crossings signs. They include, for example, the 8_{18} and 10_{123} knots that, owing to their alternating character, are not realisable with a small number of templates if these have a strictly helical geometry. Moreover, when n_T and n_S are not mutually prime they include multicomponent constructs, such as catenanes and links. In such cases, the number of components corresponds to the greatest common divisor of n_T and n_S.
For both knots and links in Fig. 4.19, the number of projected crossings, a measure of topological complexity, is equal to $n_T \cdot (n_S - 1)$. Note that this number, which does not necessarily coincide with the so-called minimal crossing number, is identical for all topologies realisable with the same number of templates and strands, (n_T, n_S). For instance, various 8-crossing knots are obtained for $(n_T = 4, n_S = 3)$. The cyclic ones, shown in Fig. 4.19, are the 8_{19} and 8_{18} knots, with C_4 symmetry, and the 3_1 and 4_1 ones, with C_2 symmetry. Non cyclic-symmetric instances are listed but not shown in Fig. 4.19 and include non-minimal representations of 5_1 and 5_2 topologies. Similarly, for $(n_T = 5, n_S = 4)$ one obtains various 15-crossing knots: four cyclic ones, including the privileged $15n_{41185}$ topology, and many more acyclic ones, such as the $12n_{242}$ knot that recurred in self-assembling simulations, see Fig. 4.16. Non-symmetric knots are listed in Appendix A.
Only links with the highest degree of cyclic symmetry are sketched. They include the molecular Solomon link [137] and Star of David [136] shown in Fig. 4.21 and respectively obtained for $(n_T = 4, n_S = 2)$ and $(n_T = 6, n_S = 2)$. New possible links are also shown for $n_T = 6$.
Some further combinations of (n_T, n_S) were explored for $n_T > 7$, as shown in Fig. 4.20.
The topology indexing scheme of Fig. 4.19 has a twofold implication.

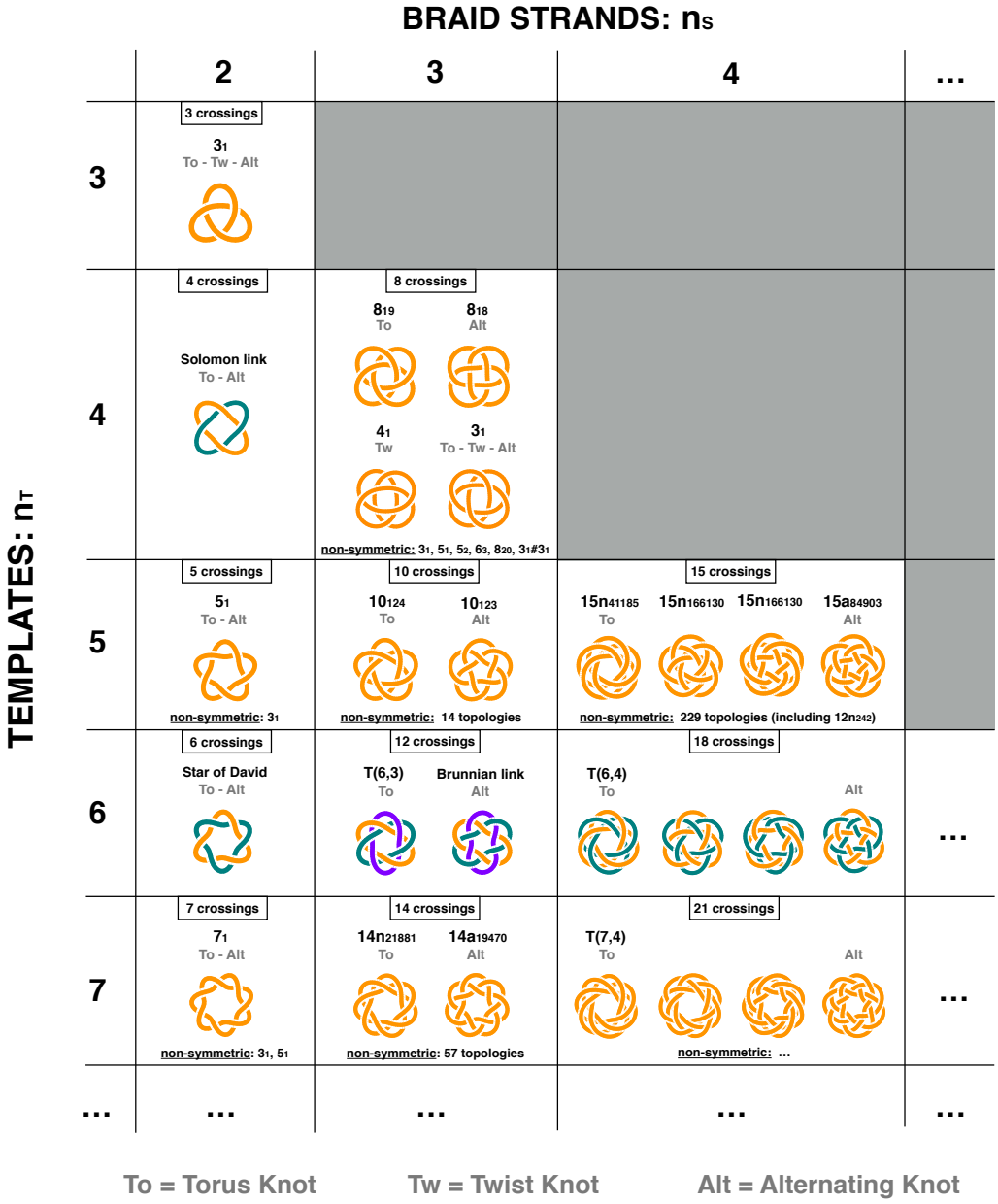


Fig. 4.19: Non-trivial knots and links with cyclic-symmetric planar representations are organised in terms of n_T and n_S. The number of projected crossings (an upper bound to the crossing number) is $n_T \cdot (n_S - 1)$ and is identical for various topologies obtainable with the same combination (n_T, n_S). For knots all inequivalent symmetric diagrams are shown, while for links only those with the highest degree of cyclic symmetry are presented for n_T=6. Various non-symmetric instances are listed too. Taken from [30].

$(\mathbf{n_T = 8, n_S = 3})$

16 projected crossings

Topology	# possible braid rep.
C_8 - symmetric	
$16a_{379778}$	2
$16n_{783154}$	2
C_4 - symmetric	
8_{18}	4
8_{19}	8
C_2 - symmetric	
3_1	48
4_1	40
8_5	48
8_{19}	16
$12n_{725}$	72
$12n_{750}$	32
$12a_{1229}$	16
$12a_{1288}$	8
$16n_{998580}$	16

$(\mathbf{n_T = 9, n_S = 2})$

9 projected crossings

Topology	# possible braid rep.
C_9 - symmetric	
9_1	2
C_3 - symmetric	
3_1	6

$(\mathbf{n_T = 10, n_S = 3})$

20 projected crossings

Topology	# possible braid rep.
C_{10} - symmetric	
T(10,3)	2
putative 20-crossing knot #1	2
C_5 - symmetric	
10_{123}	4
10_{124}	8
C_2 - symmetric	
3_1	210
8_5	80
8_{18}	60
8_{19}	120
$12a_{1210}$	80
$12a_{1229}$	20
$12a_{1288}$	10
$12n_{725}$	60
$12n_{750}$	60
$12n_{888}$	30
$16a_{377123}$	20
$16a_{377444}$	20
$16n_{783154}$	20
$16n_{998580}$	20
$16n_{1003403}$	60
putative 20-crossing knot #2	20
putative 20-crossing knot #3	10

	Dowker Code of putative 20-crossing knots for $(\mathbf{n_T = 10, n_S = 3})$
putative 20-crossing knot #1	14 16 18 20 22 24 26 28 30 32 34 36 38 40 2 4 6 8 10 12
putative 20-crossing knot #2	10 −12 14 −18 −38 26 −28 30 −6 −32 −34 −36 40 −2 4 16 −20 −22 −8 24
putative 20-crossing knot #3	10 14 −16 20 24 28 30 32 −34 6 −36 8 38 40 12 2 −4 −18 −22 26

Fig. 4.20: Symmetric topologies for the pairs $(n_T = 8, n_S = 3)$, $(n_T = 9, n_S = 2)$ and $(n_T = 10, n_S = 3)$. In the last case, putative 20-crossing knots are identified with their Dowker Code. Taken from [30].

First, it systematically recapitulates the repertoire of designable topologies in terms of two parameters that have direct bearings on the complexity of their practical realisation: the number of templates and the number of strands. The key emerging point is that ranking topologies in terms of realisation complexity, by the number of templates and braided strand, subverts the canonical order of nominal topological complexity [40]. For instance, cyclic realisations of knots with 7 minimal crossings, such as 7_1, requires more templates than prime knots

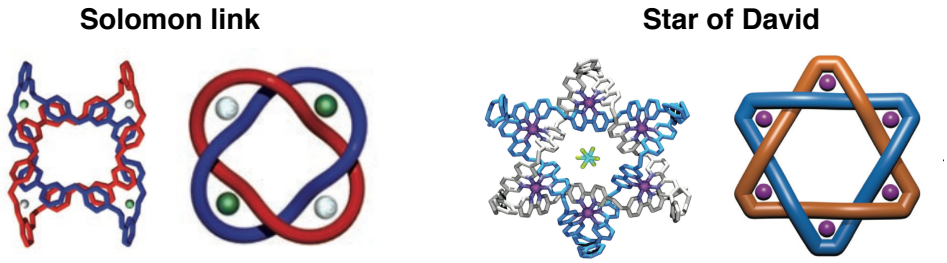


Fig. 4.21: Molecular realisation of the Solomon link [137] and the Star of David [136].

with 8 or 10 minimal crossings (e.g. 8_{19} or 10_{124}). At the same time, complex knots with 8, 10, 14 and 15 crossings can be realised with fewer templates than topologies with 6 or 7 projected crossings. This stresses the fact that the expected difficulty of realisation does not necessarily parallel the nominal complexity, and hence reinforces *a posteriori* the necessity of a systematic survey for judicious choices of target topologies [29].
Secondly, it provides a systematic route towards discovering new designable topologies by extending the range of n_S and n_T even beyond the cases considered here.
This search ought not be aimed only at the simplest types of novel knots, because complex three-dimensional geometries would be even better suited for specific goals, such as realising molecular cages [120, 144].
Arguably, S-shaped [125] or otherwise wavy building blocks ought to be more suited than constant-curvature ones (including straight and helical templates) to produce the more intricate entries in Fig. 4.19, and particularly those with the largest number of braid strands, n_S, for a given number of templates, n_T.

4.6 Summary and perspectives

In summary, we reported on a systematic scheme for discovering novel molecular topologies that are viable self-assembly targets. The method combined Monte Carlo, molecular dynamics and braiding patterns enumeration and was used to single out knot types that have the same signature features, notably cyclic symmetry and quasi-planarity, shared by all known synthetic molecular knots. It was thus established that the repertoire of admissible topologies includes only a tiny fraction of all possible knot types. In particular, there are only 6 distinct knot types that can be assembled with 5 or fewer helicate-like templates. Four of them, namely 3_1, 4_1, 5_1 and 8_{19}, have been previously obtained in a remarkable progression of synthesis strategies spanning two decades; the other two, instead, are yet to be realised experimentally. These correspond to the 10_{124} and $15n_{41185}$ knot types. With their ten and fifteen minimal crossings, respectively, these topologies surpass previously realised in terms of complexity (less than nine minimal crossings). At the same time, their characteristics of cyclic symmetry, quasi-planarity and kinetic accessibility makes them ideal targets of future molecular designing efforts.
Finally, by exhaustive enumeration of braid patterns, we found that only few

and specific types of additional knots might be addressable by using either a larger number of helical templates, or few templates but with wavy or S-shaped geometries [125], either rigid [128] or with flexible termini [27].
The simplest of these privileged topologies appear in Fig. 4.19. The shown targets include the 8_{18} topology, which was reported experimentally [132] after the completion of this study. This molecular knot was obtained with a non-minimal geometry, corresponding to $(n_t = 8, n_s = 3)$ in the scheme of Fig. 4.19, see Fig. 4.20. Its molecular representation is illustrated in Fig. 4.22. This match confirms the predictive capabilities of the enumerative scheme of Fig. 4.19.
The *in silico* exploration and optimisation of templates shapes ought to be valuable also in supramolecular DNA assembling strategies [101, 145–147] that are characterised by a good control of the local curvature of templates as well as of their spatial coordination and binding pattern.

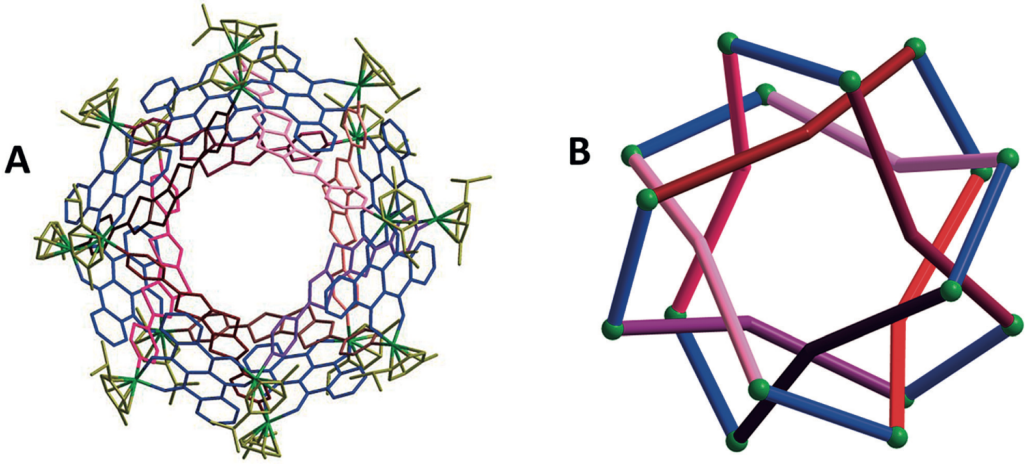


Fig. 4.22: Molecular 8_{18} knot [132] obtained with a non-minimal geometry corresponding to $(n_t = 8, n_s = 3)$.

Appendix A

All single-component knots from the exhaustive braid enumeration

The tables below list all the single-component knots (i.e. not links) obtainable with the exhaustive braid enumeration of (n_T, n_S) pairs of Fig. 4.19. The $(7, 4)$ pair is not included, due to the excessively large number of braid combinations yielding knotted patterns that exceed the complexity of tabulated knot types (available for prime components of up to 16 crossings). Knots of up to 10 crossings are denoted with the conventional Rolfsen notation. More complex knots are labelled with the Thistlethwaite notation, except for specific instances of torus knots, for which we use the conventional $T(p, q)$ notation, and very complex topologies which we fingerprint with their Dowker code.
For each (n_T, n_S) combination, the $2^{n_T(n_S-1)}$ possible braid patterns are subdivided according to several criteria. First we consider the number of crossings projected in the plane orthogonal to the axis of cyclic symmetry, which is clearly an upper bound to the minimal crossing number. Next, we separate braids that admit a closed cyclic-symmetric arrangement from those that do not. The symmetric braids are then grouped by the order of their cyclic symmetry. The number of linear braid patterns associated to a given knot type is shown in the second column.

$(\mathbf{n_T = 3, n_S = 2})$

3 projected crossings

Topology	# possible braid rep.
C_3 - symmetric	
3_1	2
Non-symmetric	
0_1	6

$(\mathbf{n_T = 4, n_S = 3})$

8 projected crossings

Topology	# possible braid rep.
C_4 - symmetric	
8_{18}	2
8_{19}	2
C_2 - symmetric	
3_1	8
4_1	4
Non-symmetric	
0_1	88
3_1	64
5_1	16
5_2	32
6_3	16
8_{20}	16
$3_1\#3_1$	8

$(\mathbf{n_T = 5, n_S = 2})$

5 projected crossings

Topology	# possible braid rep.
C_5 - symmetric	
5_1	2
Non-symmetric	
0_1	20
3_1	10

$(\mathbf{n_T = 5, n_S = 3})$

10 projected crossings

Topology	# possible braid rep.
C_5 - symmetric	
10_{123}	2
10_{124}	2
Non-symmetric	
0_1	330
3_1	200
4_1	60
5_1	60
5_2	60
6_2	80
6_3	20
8_9	20
8_{17}	20
8_{19}	20
8_{20}	20
8_{21}	60
10_{141}	20
10_{155}	10
$3_1\#3_1$	40

Fig. A.1: Symmetric and non-symmetric knots for the following (n_T, n_S) pairs: $(3,2)$, $(4,3)$, $(5,2)$, $(5,3)$. The number of linear braid patterns associated to a given knot type is shown in the second column. Taken from [30].

$(\mathbf{n_T = 5}, \mathbf{n_S = 4})$

15 projected crossings

Topology	# possible braid rep.
C_5 - symmetric	
$15a_{84903}$	2
$15n_{41185}$	2
$15n_{166130}$	4
Non-symmetric	
0_1	2910
3_1	3130
4_1	370
5_1	1480
5_2	2360
6_1	100
6_2	140
6_3	1200
7_1	200
7_2	440
7_3	520
7_4	420
7_5	500
7_6	280
7_7	20
8_7	380
8_8	660
8_{10}	680
8_{13}	320
8_{15}	180
8_{16}	140
8_{18}	80
8_{19}	310
8_{20}	1460
8_{21}	50
9_{28}	60
9_{31}	140
9_{33}	40
9_{42}	20
9_{43}	40
9_{44}	120
9_{45}	160
9_{46}	10
9_{48}	90
9_{49}	100
10_{112}	80
10_{114}	40
10_{122}	60
10_{124}	60
10_{125}	280
10_{126}	240
10_{128}	120
10_{129}	160
10_{130}	280
10_{132}	360
10_{134}	120
10_{135}	520
10_{139}	50
10_{140}	140
10_{142}	60
10_{143}	180
10_{145}	80
10_{146}	10
10_{148}	320
10_{151}	440
10_{153}	520
10_{156}	40
10_{157}	20
10_{159}	40
10_{160}	40
10_{161}=10_{162}	140
10_{164}	40
10_{165}	80
$11a_{171}$	40
$11n_{12}$	40
$11n_{23}$	80
$11n_{24}$	160
$11n_{39}$	60
$11n_{41}$	40
$11n_{46}$	80
$11n_{50}$	20
$11n_{54}$	160
$11n_{56}$	40
$11n_{61}$	120
$11n_{65}$	80
$11n_{71}$	40
$11n_{78}$	20
$11n_{82}$	40
$11n_{94}$	120
$11n_{95}$	40
$11n_{96}$	200
$11n_{98}$	40
$11n_{106}$	180
$11n_{107}$	20
$11n_{118}$	20
$11n_{132}$	40
$11n_{133}$	20
$11n_{145}$	160
$11n_{146}$	40
$11n_{147}$	40
$11n_{148}$	60
$11n_{173}$	20
$11n_{178}$	20
$11n_{179}$	20
$11n_{183}$	40
$11n_{184}$	20
$12n_{121}$	60
$12n_{242}$	20
$12n_{253}$	40
$12n_{309}$	140
$12n_{318}$	160
$12n_{323}$	200
$12n_{328}$	40
$12n_{371}$	140
$12n_{385}$	80
$12n_{425}$	60
$12n_{426}$	20
$12n_{439}$	100
$12n_{443}$	80
$12n_{451}$	80
$12n_{488}$	100
$12n_{548}$	40
$12n_{591}$	40
$12n_{646}$	40

$12n_{702}$	80
$12n_{725}$	10
$12n_{730}$	40
$12n_{749}$	60
$12n_{811}$	20
$12n_{829}$	20
$12n_{835}$	40
$12n_{868}$	10
$13n_{225}$	40
$13n_{288}$	80
$13n_{501}$	60
$13n_{519}$	20
$13n_{584}$	40
$13n_{586}$	20
$13n_{592}$	120
$13n_{601}$	40
$13n_{603}$	40
$13n_{606}$	120
$13n_{608}$	80
$13n_{1192}$	20
$13n_{1644}$	40
$13n_{1692}$	80
$13n_{1716}$	40
$13n_{1718}$	40
$13n_{1719}$	80
$13n_{1724}$	80
$13n_{1727}$	40
$13n_{1734}$	20
$13n_{1735}$	40
$13n_{1739}$	40
$13n_{1931}$	120
$13n_{1945}$	40
$13n_{1957}$	20
$13n_{2303}$	40
$13n_{2436}$	40
$13n_{2442}$	40
$13n_{2491}$	40
$13n_{2787}$	120
$13n_{3023}$	40
$13n_{3351}$	20

$13n_{3393}$	80
$13n_{3582}$	40
$13n_{3611}$	40
$13n_{3956}$	40
$13n_{3958}$	40
$13n_{3969}$	40
$13n_{3973}$	40
$13n_{3978}$	20
$13n_{3979}$	40
$13n_{3982}$	40
$13n_{3998}$	40
$13n_{4003}$	40
$13n_{4035}$	60
$13n_{4051}$	10
$13n_{4079}$	60
$13n_{4080}$	60
$13n_{4634}$	20
$13n_{5018}$	10
$14n_{6174}$	10
$14n_{21472}$	20
$14n_{22172}$	20
$14n_{22583}$	20
$14n_{23344}$	20
$15n_{40180}$	20
$15n_{40184}$	40
$15n_{40185}$	20
$15n_{41127}$	20
$15n_{41189}$	20
$15n_{41193}$	60
$15n_{41223}$	20
$15n_{41235}$	40
$15n_{43517}$	30
$15n_{45460}$	10
$15n_{46935}$	40
$15n_{46936}$	20
$15n_{47800}$	40
$15n_{48957}$	20
$15n_{49035}$	40
$15n_{51709}$	20
$15n_{52941}$	10

$15n_{52944}$	20
$15n_{53947}$	20
$15n_{53948}$	20
$15n_{56026}$	20
$15n_{59004}$	40
$15n_{59005}$	20
$15n_{59007}$	20
$15n_{71113}$	20
$15n_{107628}$	40
$15n_{124826}$	40
$15n_{125031}$	40
$15n_{125991}$	20
$15n_{126002}$	20
$15n_{126008}$	20
$15n_{126010}$	40
$15n_{126011}$	20
$15n_{126024}$	10
$15n_{127000}$	20
$15n_{127094}$	10
$15n_{127330}$	10
$15n_{127609}$	20
$15n_{127610}$	20
$15n_{127630}$	40
$15n_{127654}$	40
$15n_{163844}$	20
$15n_{163860}$	20
$15n_{166131}$	10
$15n_{166806}$	10
$3_1\#3_1$	730
$3_1\#4_1$	60
$3_1\#5_1$	280
$3_1\#5_2$	400
$3_1\#6_3$	80
$3_1\#8_{20}$	60
$3_1\#8_{19}$	40
$3_1\#3_1\#3_1$	10

Fig. A.2: Symmetric and non-symmetric knots for the following (n_T, n_S) pairs: $(5, 4)$. The number of linear braid patterns associated to a given knot type is shown in the second column. Taken from [30].

$(\mathbf{n_T = 7, n_S = 2})$

7 projected crossings

Topology	# possible braid rep.
C7 - symmetric	
7_1	2
Non-symmetric	
0_1	70
3_1	42
5_1	14

$(\mathbf{n_T = 7, n_S = 3})$

14 projected crossings

Topology	# possible braid rep.
C7 - symmetric	
$14a_{19470}$	2
$14n_{21881}$	2
Non-symmetric	
0_1	2688
3_1	2884
4_1	224
5_1	1260
5_2	1400
6_2	168
6_3	896
7_1	280
7_3	280
7_5	560
8_7	336
8_9	28
8_{10}	336
8_{16}	168
8_{17}	28
8_{18}	112
8_{19}	280
8_{20}	868
8_{21}	140
10_{17}	28
10_{48}	56
10_{79}	28
10_{91}	56
10_{99}	14
10_{104}	28
10_{109}	28
10_{112}	168
10_{118}	28
10_{124}	70
10_{125}	196
10_{126}	140
10_{139}	112
10_{141}	28
10_{143}	280
10_{148}	280
10_{155}	14
10_{157}	70
10_{159}	140
10_{161}=10_{162}	168
$12a_{819}$	28
$12a_{1209}$	28
$12a_{1211}$	28
$12n_{242}$	28
$12n_{468}$	56
$12n_{675}$	56
$12n_{708}$	56
$12n_{709}$	56
$12n_{721}$	28
$12n_{749}$	112
$12n_{751}$	28
$12n_{829}$	140
$14n_{21882}$	28
$14n_{24169}$	28
$14n_{27039}$	28
$14n_{27120}$	14
$3_1\#3_1$	588
$3_1\#5_1$	168
$5_1\#5_1$	14

Fig. A.3: Symmetric and non-symmetric knots for the following (n_T, n_S) pairs: $(7, 2)$, $(7, 3)$. The number of linear braid patterns associated to a given knot type is shown in the second column. Taken from [30].

Bibliography

[1] Delbruck, M. *Mathematical Problems in the Biological Sciences* (American Mathematical Society, 1962).

[2] Sumners, D. W. & Whittington, S. G. Knots in self-avoiding walks. *Journal of Physics A: Mathematical and General* **21**, 1689 (1988).

[3] Micheletti, C., Marenduzzo, D. & Orlandini, E. Polymers with spatial or topological constraints: Theoretical and computational results. *Physics Reports* **504**, 1 – 73 (2011).

[4] Micheletti, C. & Orlandini, E. Numerical study of linear and circular model DNA chains confined in a slit: Metric and topological properties. *Macromolecules* **45**, 2113 – 2121 (2012).

[5] Rieger, F. C. & Virnau, P. A monte carlo study of knots in long double-stranded DNA chains. *PLOS Computational Biology* **12**, 1 – 11 (2016).

[6] Wasserman, S. & Cozzarelli, N. Biochemical topology: applications to DNA recombination and replication. *Science* **232**, 951 – 960 (1986).

[7] Liu, Z., Deibler, R. W., Chan, H. S. & Zechiedrich, L. The why and how of DNA unlinking. *Nucleic Acids Research* **37**, 661 – 671 (2009).

[8] Sumners, D. W. Untangling DNA. *Mathematics Intelligencer* **12**, 71 – 80 (1990).

[9] Ashley, C. & Lee, J. S. A triplex-mediated knot between separated polypurine-polypyrimidine tracts in circular DNA blocks transcription by escherichia coli RNA polymerase. *DNA and Cell Biology* **19**, 235 – 241 (2000).

[10] Portugal, J. & Rodriguez-Campos, A. T7 RNA polymerase cannot transcribe through a highly knotted DNA template. *Nucleic Acids Research* **24**, 4890 – 4894 (1996).

[11] Deibler, R. W., Mann, J. K., Sumners, D. W. L. & Zechiedrich, L. Hin-mediated DNA knotting and recombining promote replicon dysfunction and mutation. *BMC Molecular Biology* **8**, 44 (2007).

[12] Vologodskii, A. V. *et al.* Mechanism of topology simplification by Type II DNA topoisomerases. *Proceedings of the National Academy of Sciences* **98**, 3045 – 3049 (2001).

[13] Wang, J. C. DNA topoisomerases. *Annual Review of Biochemistry* **65**, 635 – 692 (1996).

[14] Rybenkov, V. V., Ullsperger, C., Vologodskii, A. V. & Cozzarelli, N. R. Simplification of DNA topology below equilibrium values by Type II Topoisomerases. *Science* **277**, 690 – 693 (1997).

[15] Sumners, D. W., Ernst, C., Spengler, S. J. & Cozzarelli, N. R. Analysis of the mechanism of DNA recombination using tangles. *Quarterly Reviews of Biophysics* **28**, 253 – 313 (1995).

[16] Liu, Z., Mann, J. K., Zechiedrich, E. L. & Chan, H. S. Topological information embodied in local juxtaposition geometry provides a statistical mechanical basis for unknotting by Type-2 DNA Topoisomerases. *Journal of Molecular Biology* **361**, 268 – 285 (2006).

[17] Liu, Z., Zechiedrich, L. & Chan, H. S. Action at hooked or twistedhooked DNA juxtapositions rationalizes unlinking preference of Type-2 Topoisomerases. *Journal of Molecular Biology* **400**, 963 – 982 (2010).

[18] Stasiak, A., Katritch, V., Bednar, J., Michoud, D. & Dubochet, J. Electrophoretic mobility of DNA knots. *Nature* **384**, 122 (1996).

[19] Trigueros, S., Arsuaga, J., Vazquez, M. E., Sumners, D. W. & Roca, J. Novel display of knotted DNA molecules by two-dimensional gel electrophoresis. *Nucleic Acids Research* **29**, e67 (2001).

[20] Arsuaga, J. *et al.* DNA knots reveal a chiral organization of DNA in phage capsids. *Proceedings of the National Academy of Sciences* **102**, 9165 – 9169 (2005).

[21] Marenda, M., Orlandini, E. & Micheletti, C. Sorting ring polymers by knot type with modulated nanochannels. *Soft Matter* **13**, 795 – 802 (2017).

[22] Virnau, P., Mirny, L. A. & Kardar, M. Intricate knots in proteins: Function and evolution. *PLOS Computational Biology* **2**, 1 – 6 (2006).

[23] Virnau, P., Mallam, A. & Jackson, S. Structures and folding pathways of topologically knotted proteins. *Journal of Physics: Condensed Matter* **23**, 033101 (2010).

[24] Potestio, R., Micheletti, C. & Orland, H. Knotted vs. unknotted proteins: Evidence of knot-promoting loops. *PLOS Computational Biology* **6**, 1 – 10 (2010).

[25] Lou, S.-C. *et al.* The knotted protein UCH-L1 exhibits partially unfolded forms under native conditions that share common structural features with its kinetic folding intermediates. *Journal of Molecular Biology* **428**, 2507 – 2520 (2016).

[26] Ayme, J.-F., Beves, J. E., Campbell, C. J. & Leigh, D. A. Template synthesis of molecular knots. *Chemical Society Reviews* **42**, 1700 – 1712 (2013).

[27] Ponnuswamy, N., Cougnon, F. B. L., Dan Pantos, G. & Sanders, J. K. M. Homochiral and meso figure eight knots and a solomon link. *Journal of the American Chemical Society* **136**, 8243 – 8251 (2014).

[28] Barnes, J. C. & Mirkin, C. A. Profile of Jean-Pierre Sauvage, Sir J. Fraser Stoddart, and Bernard L. Feringa, 2016 Nobel Laureates in Chemistry. *Proceedings of the National Academy of Sciences* **114**, 620 – 625 (2017).

[29] Chambron, J.-C. & Sauvage, J.-P. Topologically complex molecules obtained by transition metal templation: it is the presentation that determines the synthesis strategy. *New Journal of Chemistry* **37**, 49 – 57 (2013).

[30] Marenda, M., Orlandini, E. & Micheletti, C. Discovering privileged topologies of molecular knots with self-assembling models. *Nature Communications* **9**, 3051 (2018).

[31] Adams, C. C. *The knot book: an elementary introduction to the mathematical theory of kntos* (The American Mathematical Society, 2004).

[32] Cromwell, P. *Knots and Links* (Cambridge University Press, 1994).

[33] Tubiana, L., Polles, G., Orlandini, E. & Micheletti, C. KymoKnot: A web server and software package to identify and locate knots in trajectories of linear or circular polymers. *The European Physical Journal E* **41**, 72 (2018).

[34] Sossinsky, A. *Knots: mathematic with a twist* (Harvard University Press, 2002).

[35] Scharein, R. Knotplot. *http://www.knotplot.com/* .

[36] Fenlon, E. E. Open problems in chemical topology. *European Journal of Organic Chemistry* 5023 – 5035 (2008).

[37] Ernst, C. & Sumners, D. W. The growth of the number of prime knots. *Mathematical Proceedings of the Cambridge Philosophical Society* **102**, 303 – 315 (1987).

[38] Hoste, J., Thistlethwaite, M. & Weeks, J. The first 1,701,936 knots. *Mathematical Intelligencer* **20**, 33 – 48 (1998).

[39] Alexander, J. W. & Briggs, G. B. On types of knotted curves. *Annals of Mathematics* **28**, 562 – 586 (1926).

[40] Rolfsen, D. *Knots and links* (The American Mathematical Society, 1976).

[41] Alexander, J. W. Topological invariants of knots and links. *Transactions of the American Mathematical Society* **30**, 275 – 306 (1928).

[42] Hoste, J. & Thistlethwaite, M. Knotscape 1.01. *http://www.math.utk.edu/~morwen/knotscape.html* .

[43] Virnau, P., Kantor, Y. & Kardar, M. Knots in globule and coil phases of a model polyethylene. *The Journal of American Chemical Society* **127** (2005).

[44] Tubiana, L., Orlandini, E. & Micheletti, C. Probing the entanglement and locating knots in ring polymers: A comparative study of different arc closure schemes. *Progress of Theoretical Physics Supplement* **191**, 192 – 204 (2011).

[45] Mansfield, M. L. & Douglas, J. F. Properties of knotted ring polymers. I. Equilibrium dimensions. *The Journal of Chemical Physics* **133**, 044903 (2010).

[46] Alexander, J. W. A lemma on systems of knotted curves. *Proceedings of the National Academy of Sciences* **9**, 93 – 95 (1923).

[47] Rubinstein, M. & Colby, R. H. *Polymer Physics* (Oxford Univerisity Press, 2003).

[48] Kremer, K. & Grest, G. S. Dynamics of entangled linear polymer melts: a moleculardynamics simulation. *The Journal of Chemical Physics* **92**, 5057 – 5086 (1990).

[49] Suma, A., Orlandini, E. & Micheletti, C. Knotting dynamics of DNA chains of different length confined in nanochannels. *Journal of Physics: Condensed Matter* **27**, 354102 (2015).

[50] Peliti, L. *Statistical mechanics in a nutshell* (Princeton University Press, 2011).

[51] Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics* **117**, 1 – 19 (1995).

[52] Trefz, B., Siebert, J. & Virnau, P. How molecular knots can pass through each other. *Proceedings of the National Academy of Sciences* **111**, 7948 – 7951 (2014).

[53] Saitta, A. M., Soper, P. D., Wasserman, E. & Klein, M. L. Influence of a knot on the strength of a polymer strand. *Nature* **399**, 46 – 48 (1999).

[54] Farago, O., Kantor, Y. & Kardar, M. Pulling knotted polymers. *Europhysics Letters* **60**, 53 (2002).

[55] Caraglio, M., Micheletti, C. & Orlandini, E. Stretching response of knotted and unknotted polymer chains. *Physical Review Letters* **115**, 188301 (2015).

[56] Soh, B. W., Narsimhan, V., Klotz, A. R. & Doyle, P. S. Knots modify the coil-stretch transition in linear DNA polymers. *Soft Matter* **14**, 1689 – 1698 (2018).

[57] Rosa, A., Di Ventra, M. & Micheletti, C. Topological jamming of spontaneously knotted polyelectrolyte chains driven through a nanopore. *Physical Review Letters* **109**, 118301 (2012).

[58] Suma, A., Rosa, A. & Micheletti, C. Pore translocation of knotted polymer chains: How friction depends on knot complexity. *ACS Macro Letters* **4**, 1420–1424 (2015).

[59] Narsimhan, V., Renner, C. B. & Doyle, P. S. Translocation dynamics of knotted polymers under a constant or periodic external field. *Soft Matter* **12**, 5041 – 5049 (2016).

[60] Virnau, P., Rieger, F. C. & Reith, D. Influence of chain stiffness on knottedness in single polymers. *Biochemical Society Transactions* **41**, 528 – 532 (2013).

[61] Plesa, C. *et al.* Direct observation of DNA knots using a solid-state nanopore. *Nature Nanotechnology* **11**, 1093 – 1097 (2016).

[62] Suma, A. & Micheletti, C. Pore translocation of knotted DNA rings. *Proceedings of the National Academy of Sciences* **114**, E2991 – E2997 (2017).

[63] Hsieh, T.-S. Knotting of the circular duplex DNA by Type II DNA topoisomerase from Drosophila melanogaster. *Journal of Biological Chemistry* **258**, 8413 – 8420 (1983).

[64] Viovy, J.-L. Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms. *Reviews of Modern Physics* **72**, 813 – 872 (2000).

[65] Slater, G. W. *et al.* Theory of DNA electrophoresis: A look at some current challenges. *Electrophoresis* **21**, 3873 – 3887 (2000).

[66] Levene, S. Analysis of DNA topoisomers, knots, and catenanes by agarose gel electrophoresis. *Methods in molecular biology (Clifton, N.J.)* **582**, 11 – 25 (2009).

[67] Weber, C., Stasiak, A., Rios, P. D. L. & Dietler, G. Numerical simulation of gel electrophoresis of DNA knots in weak and strong electric fields. *Biophysical Journal* **90**, 3100 – 3105 (2006).

[68] Michieletto, D., Marenduzzo, D. & Orlandini, E. Topological patterns in two-dimensional gel electrophoresis of DNA knots. *Proceedings of the National Academy of Sciences* **112**, E5471 – E5477 (2015).

[69] Liu, L. F., Depew, R. E. & Wang, J. C. Knotted single-stranded DNA rings: A novel topological isomer of circular single-stranded DNA formed by treatment with escherichia coli ω protein. *Journal of Molecular Biology* **106**, 439 – 452 (1976).

[70] Liu, L. F., Perkocha, L., Calendar, R. & Wang, J. C. Knotted DNA from bacteriophage capsids. *Proceedings of the National Academy of Sciences* **78**, 5498 – 5502 (1981).

[71] Liu, L. F., Davis, J. L. & Calendar, R. Novel topologically knotted DNA from bacteriophage P4 capsids: studies with DNA topoisomerases. *Nucleic Acids Research* **9**, 3979 – 3989 (1981).

[72] Wang, J. C., Martin, K. V. & Calendar, R. Sequence similarity of the cohesive ends of coliphage P4, P2, and 186 deoxyribonucleic acid. *Biochemistry* **12**, 2119 – 2123 (1973).

[73] Isaksen, M., Julien, B., Calendar, R. & Lindqvist, B. H. Isolation of knotted DNA from coliphage P4. *Methods in Molecular Biology* **94**, 69 – 74 (1999).

[74] Arsuaga, J., Vazquez, M., Trigueros, S., Sumners, D. W. & Roca, J. Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proceedings of the National Academy of Sciences* **99**, 5373 – 5377 (2002).

[75] Chou, C.-F. *et al.* Sorting by diffusion: An asymmetric obstacle course for continuous molecular separation. *Proceedings of the National Academy of Sciences* **96**, 13762 – 13765 (1999).

[76] Dorfman, K. D. DNA electrophoresis in microfabricated devices. *Reviews of Modern Physics* **82**, 2903 – 2947 (2010).

[77] Dorfman, K. D., King, S. B., Olson, D. W., Thomas, J. D. P. & Tree, D. R. Beyond gel electrophoresis: Microfluidic separations, fluorescence burst analysis, and DNA stretching. *Chemical Reviews* **113**, 2584 – 2667 (2013).

[78] Doyle, P. S., Bibette, J., Bancaud, A. & Viovy, J.-L. Self-assembled magnetic matrices for DNA separation chips. *Science* **295**, 2237 – 2237 (2002).

[79] Volkmuth, W. D. & Austin, R. H. DNA electrophoresis in microlithographic arrays. *Nature* **358**, 600 – 602 (1992).

[80] Fanzio, P. *et al.* DNA detection with a polymeric nanochannel device. *Lab on a Chip* **11**, 2961 – 2966 (2011).

[81] Fu, J., Schoch, R. B., Stevens, A. L., Tannenbaum, S. R. & Han, J. A patterned anisotropic nanofluidic sieving structure for continuous-flow separation of DNA and proteins. *Nature Nanotechnology* **2**, 121 – 128 (2007).

[82] Fu, J., Yoo, J. & Han, J. Molecular sieving in periodic free-energy landscapes created by patterned nanofilter arrays. *Physical Review Letters* **97**, 018103 (2006).

[83] Rahong, S. *et al.* Three-dimensional nanowire structures for ultra-fast separation of DNA, protein and RNA molecules. *Scientific Reports* **5**, 10584 (2015).

[84] Frykholm, K. *et al.* Fast size-determination of intact bacterial plasmids using nanofluidic channels. *Lab on a Chip* **15**, 2739 – 2743 (2015).

[85] Ollila, S. T. T., Denniston, C., Karttunen, M. & Ala-Nissila, T. Biopolymer filtration in corrugated nanochannels. *Physical Review Letters* **112**, 118301 (2014).

[86] Persson, F., Utko, P., Reisner, W., Larsen, N. B. & Kristensen, A. Confinement spectroscopy: Probing single DNA molecules with tapered nanochannels. *Nano Letters* **9**, 1382 – 1385 (2009).

[87] Fanzio, P. *et al.* Modulating DNA translocation by a controlled deformation of a PDMS nanochannel device. *Scientific Reports* **2**, 791 (2012).

[88] Micheletti, C. & Orlandini, E. Knotting and metric scaling properties of DNA confined in nano-channels: a Monte Carlo study. *Soft Matter* **8**, 10959 – 10968 (2012).

[89] Orlandini, E. & Micheletti, C. Knotting of linear DNA in nano-slits and nano-channels: a numerical study. *Journal of Biological Physics* **39**, 267 – 275 (2013).

[90] Micheletti, C. & Orlandini, E. Knotting and unknotting dynamics of DNA strands in nanochannels. *ACS Macro Letters* **3**, 876 – 880 (2014).

[91] Jacobs, M. H. *Diffusion Processes* (Springer Berlin Heidelberg, 1967).

[92] Zwanzig, R. Diffusion past an entropy barrier. *The Journal of Physical Chemistry* **96**, 3926 – 3930 (1992).

[93] Lifson, S. & Jackson, J. L. On the self-diffusion of ions in a polyelectrolyte solution. *The Journal of Chemical Physics* **36**, 2410 – 2414 (1962).

[94] Hanggi, P., Talkner, P. & Borkovec, M. Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics* **62**, 251 – 341 (1990).

[95] Reimann, P. *et al.* Diffusion in tilted periodic potentials: Enhancement, universality, and scaling. *Physical Review E* **65**, 031104 (2002).

[96] Burada, P. S., Hnggi, P., Marchesoni, F., Schmid, G. & Talkner, P. Diffusion in confined geometries. *ChemPhysChem* **10**, 45 – 54 (2009).

[97] Weiss, L. B., Marenda, M., Micheletti, C. & Likos, C. N. Translocation of polymers of various knot types through modulated channels with MPC. *Work in preparation* (2018).

[98] Malevanets, A. & Kapral, R. Mesoscopic model for solvent dynamics. *The Journal of Chemical Physics* **110**, 8605 – 8613 (1999).

[99] Ercolini, E. *et al.* Fractal dimension and localization of DNA knots. *Physical Review Letters* **98**, 058102 (2007).

[100] Mikkelsen, M., Reisner, W., Flyvbjerg, H. & Kristensen, A. Pressure-driven DNA in nanogroove arrays: Complex dynamics leads to length- and topology-dependent separation. *Nano Letters* **11**, 1598 – 1602 (2011).

[101] Aldaye, F. A., Palmer, A. L. & Sleiman, H. F. Assembling materials with DNA as the guide. *Science* **321**, 1795 – 1799 (2008).

[102] Sciortino, F., Giacometti, A. & Pastore, G. Phase diagram of janus particles. *Physical Reviews Letters* **103**, 237801 (2009).

[103] Romano, F. & Sciortino, F. Patchy from the bottom up. *Nature Materials* **10**, 171 (2011).

[104] Romano, F. & Sciortino, F. Patterning symmetry in the rational design of colloidal crystals. *Nature Communications* **3:975** (2012).

[105] Damasceno, P. F., Engel, M. & Glotzer, S. C. Predictive self-assembly of polyhedra into complex structures. *Science* **337**, 453 – 457 (2012).

[106] Furst, E. M. Directed self-assembly. *Soft Matter* **9**, 9039 – 9045 (2013).

[107] Mallory, S. A., Valeriani, C. & Cacciuto, A. An active approach to colloidal self-assembly. *Annual Review of Physical Chemistry* **69**, 59 – 79 (2017).

[108] Sauvage, J.-P. & Dietrich-Buchecker, C. *Catenanes, Rotaxanes and Knots* (VCH-Wiley, Weinheim, Germany, 1999).

[109] Gil-Ramirez, G., Leigh, D. A. & Stephens, A. J. Catenanes: Fifty years of molecular links. *Angewandte Chemie International Edition* **54**, 6110 – 6150 (2015).

[110] Marcos, V. *et al.* Allosteric initiation and regulation of catalysis with a molecular knot. *Science* **352**, 1555 – 1559 (2016).

[111] Ayme, J.-F. *et al.* Strong and selective anion binding within the central cavity of molecular knots and links. *Journal of the American Chemical Society* **137**, 9812 – 9815 (2015).

[112] Bilbeisi, R. A. *et al.* [C-H $\cdots$ anion] interactions mediate the templation and anion binding properties of topologically non-trivial metal-organic structures in aqueous solutions. *Chemical Science* **7**, 2524 – 2531 (2016).

[113] Chichak, K. S., Peters, A. J., Cantrill, S. J. & Stoddart, J. F. Nanoscale borromeates. *The Journal of Organic Chemistry* **70**, 7956 – 7962 (2005).

[114] Yates, C. R., Benitez, D., Khan, S. I. & Stoddart, J. F. Hexafunctionalized borromeates using olefin cross metathesis. *Organic Letters* **9**, 2433 – 2436 (2007).

[115] Lukin, O. & Voegtle, F. Knotting and threading of molecules: Chemistry and chirality of molecular knots and their assemblies. *Angewandte Chemie International Edition* **44**, 1456 – 1477 (2005).

[116] Horner, K. E., Miller, M. A., Steed, J. W. & Sutcliffe, P. M. Knot theory in modern chemistry. *Chemical Society Reviews* **45**, 6432 – 6448 (2016).

[117] Liu, Y. *et al.* Weaving of organic threads into a crystalline covalent organic framework. *Science* **351**, 365 – 369 (2016).

[118] Lewandowska, U. *et al.* A triaxial supramolecular weave. *Nature Chemistry* **9**, 1068 – 1072 (2017).

[119] Wang, Z. *et al.* Molecular weaving via surface-templated epitaxy of crystalline coordination networks. *Nature Communications* **8**, 14442 (2017).

[120] Riddell, I. A. *et al.* Cation- and anion-exchanges induce multiple distinct rearrangements within metallosupramolecular architectures. *Journal of the American Chemical Society* **136**, 9491 – 9498 (2014).

[121] Cougnon, F. B. L. Engineering protein self-assembly: A new approach for the design of octahedral cages. *ChemBioChem* **17**, 2296 – 2298 (2016).

[122] Balzani, V., Credi, A., Raymo, F. M. & Stoddart, J. F. Artificial molecular machines. *Angewandte Chemie International Edition* **39**, 3348 – 3391 (2000).

[123] Prakasam, T. *et al.* Simultaneous self-assembly of a [2]catenane, a trefoil knot, and a solomon link from a simple pair of ligands. *Angewandte Chemie International Edition* **52**, 9956 – 9960 (2013).

[124] Gil-Ramirez, G. *et al.* Tying a molecular overhand knot of single handedness and asymmetric catalysis with the corresponding pseudo-D3-symmetric trefoil knot. *Journal of the American Chemical Society* **138**, 13159 – 13162 (2016).

[125] Beves, J. E., Danon, J. J., Leigh, D. A., Lemonnier, J.-F. & Vitorica-Yrezabal, I. J. A solomon link through an interwoven molecular grid. *Angewandte Chemie International Edition* **54**, 7555 – 7559 (2015).

[126] Cougnon, F. B. L., Ponnuswamy, N., Pantos, G. D. & Sanders, J. K. M. Molecular motion of donor-acceptor catenanes in water. *Organic and Biomolecular Chemistry* **13**, 2927 – 2930 (2015).

[127] Meyer, C. D. *et al.* The dynamic chemistry of molecular borromean rings and solomon knots. *Chemistry - A European Journal* **16**, 12570 – 12581 (2010).

[128] Schouwey, C. *et al.* Self-assembly of a giant molecular solomon link from 30 subcomponents. *Angewandte Chemie International Edition* **53**, 11261 – 11265 (2014).

[129] Ayme, J.-F. *et al.* Lanthanide template synthesis of a molecular trefoil knot. *Journal of the American Chemical Society* **136**, 13142 – 13145 (2014).

[130] Ayme, J.-F. *et al.* A synthetic molecular pentafoil knot. *Nature Chemistry* **4**, 15 – 20 (2011).

[131] Danon, J. J. *et al.* Braiding a molecular knot with eight crossings. *Science* **355**, 159 – 162 (2017).

[132] Hwan, K. D. *et al.* Coordination-driven self-assembly of a molecular knot comprising sixteen crossings. *Angewandte Chemie International Edition* **57**, 5669 – 5673 (2018).

[133] Dietrich-Buchecker, C. O. & Sauvage, J. P. A synthetic molecular trefoil knot. *Angewandte Chemie International Edition in English* **28**, 189 – 192 (1989).

[134] Chichak, K. S. *et al.* Molecular borromean rings. *Science* **304**, 1308 – 1312 (2004).

[135] Nierengarten, J. F., Dietrich-Buchecker, C. O. & Sauvage, J. P. Synthesis of a doubly interlocked [2]-catenane. *Journal of the American Chemical Society* **116**, 375 – 376 (1994).

[136] Leigh, D. A., Pritchard, R. G. & Stephens, A. J. A Star of David catenane. *Nature Chemistry* **6**, 978 – 982 (2014).

[137] Pentecost, C. D. *et al.* A molecular solomon link. *Angewandte Chemie International Edition* **46**, 218 – 222 (2007).

[138] Polles, G., Marenduzzo, D., Orlandini, E. & Micheletti, C. Self-assembling knots of controlled topology by designing the geometry of patchy templates. *Nature Communications* **6**, 6423 (2015).

[139] Polles, G., Orlandini, E. & Micheletti, C. Optimal self-assembly of linked constructs and catenanes via spatial confinement. *ACS Macro Letters* **5**, 931 – 935 (2016).

[140] Dietrich-Buchecker, C. O., Sauvage, J. P. & Kern, J. M. Templated synthesis of interlocked macrocyclic ligands: the catenands. *Journal of the American Chemical Society* **106**, 3043 – 3045 (1984).

[141] Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A* **34**, 827 – 828 (1978).

[142] Miller, M. A. & Wales, D. J. Novel structural motifs in clusters of dipolar spheres: knots, links, and coils. *The Journal of Physical Chemistry B* **109**, 23109 – 23112 (2005).

[143] Forgan, R. S., Sauvage, J.-P. & Stoddart, J. F. Chemical topology: Complex molecular knots, links, and entanglements. *Chemical Reviews* **111**, 5434 – 5464 (2011).

[144] Leenders, S. H. A. M., Gramage-Doria, R., de Bruin, B. & Reek, J. N. H. Transition metal catalysis in confined spaces. *Chemical Society Reviews* **44**, 433 – 448 (2015).

[145] Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297 – 302 (2006).

[146] Han, D., Pal, S., Liu, Y. & Yan, H. Folding and cutting DNA into reconfigurable topological nanostructures. *Nature Nanotechnology* **5**, 712 – 717 (2010).

[147] Zhang, F., Nangreave, J., Liu, Y. & Yan, H. Structural DNA nanotechnology: State of the art and future perspective. *Journal of the American Chemical Society* **136**, 11198 – 11211 (2014).